Global navigation satellite system (GNSS) signal simulator: an analysis of the effects of the local environment and atmosphere on receiver positioning

Smith, Andrew

Award date:
2007

Awarding institution:
University of Bath

GLOBAL NAVIGATION SATELLITE SYSTEM SIGNAL SIMULATOR

AN ANALYSIS OF THE EFFECTS OF THE LOCAL ENVIRONMENT AND ATMOSPHERE ON RECEIVER POSITIONING

Submitted by Andrew M. Smith
for the degree of
Doctor of Philosophy

University of Bath
Department of Electronic and Electrical Engineering

October 2007

Abstract

Global Navigation Satellite Systems can provide position, velocity and time information to users using receiver hardware. The United States developed Global Positioning System (GPS) is the only current fully operational system; however further systems are in development. The GPS has shown considerable success for navigation, but it still has a number of problems that limit its accuracy. The two main problems are the ionosphere and local environment of the receiver. The ionosphere causes a delay and random rapid shifts in phase and amplitude (scintillation) to the signal. The local environment can provide the signal with multiple routes (multi-path) to the receiver. In this project a GPS signal simulator is developed, which models the effects of the ionosphere and multi-path on the modulated signals. The focus is made on the GPS system as the simulator measurements can be compared to the real measurements; however other systems will be considered in the future. A number of experiments investigating multi-path and ionospheric effects on a receiver's ability to track the signals have been completed. The simulator has been used to replicate a real local multi-path environment and the results have been compared. Further investigations of the multi-path have shown a unique multi-path signature in the receiver power output. The later part of the thesis describes a case study investigating a short but rapid period of scintillation observed on three receivers based in Norway. An analysis of the multi-path environment was completed, but was found not to be the cause. The ionosphere was investigated using equipment based across Scandinavia. The equipment showed that geomagnetic conditions were disturbed at the time of the event. The GPS measurements were compared with all-sky camera data to show that the scintillation can be attributed to the GPS signal path crossing electron density structures associated with the aurora.

Acknowledgements

I would like to acknowledge the support provided by the UK Engineering and Physical Sciences Research Council (EPSRC), and by my industrial sponsor BAE SYSTEMS. A thank you to my supervisors, Dr Cathryn Mitchell and Dr Robert Watson for their support, advice and help throughout the project. A thank you to my girlfriend Michaela for her support and encouragement.

An acknowledgement to the International Reference Ionosphere group for their freely available ionosphere model and Northwest Research Associates Inc. for providing an academic license for the WBMOD scintillation model. I would also like to acknowledge the Finnish Meteorological Institute (FMI) for access to their magnetometer data and help with the interpretation.

A final thank you to my Mother for my inquisitive mind; an essential requirement for this project.

Contents

List of Figures

List of Tables

Abbreviations

GPS Global Positioning System

GNSS Global Navigation Satellite System

GLONASS Globalnaya Navigatsionnay Sputnikovaya Sistema

GEE Ground Electronic Equipment

DFT Discrete Fourier Transform

FFT Fast Fourier Transform

IFFT Inverse Fast Fourier Transform

LORAN LOng RAnge Navigation system

CDMA Code Division Multiple Access

DSCDMA Direct Sequence CDMA

FHCDMA Frequency Hopping CDMA

TDMA Time Division Multiple Access

FDMA Frequency Division Multiple Access

CDM Code Division Multiplexing

DSCDM Direct Sequence CDM

FHCDM Frequency Hopping CDM

TDM Time Division Multiplexing

FDM Frequency Division Multiplexing

NAVSAT Navy Navigation Satellite System

C/A Coarse Acquisition

DoD United States Department of Defence

SPS Standard Positioning Service

PPS Precise Positioning Service

P Precise

RF Radio Frequency

IF Intermediate Frequency

PLL Phase Locked Loop

FLL Frequency Locked Loop

DLL Delay Locked Loop

VCO Voltage Controlled Oscillator

BPSK Binary Phase Shift Keying

QPSK Quadrature Phase Shift Keying

IIR Infinite Impulse Response

FIR Finite Impulse Response

AC Alternating Current

DC Direct Current

SCC System Control Centre

TTCC Telemetry, Tracking and Control Centre

TEC Total Electron Content

IRI International Reference Ionosphere model

GDOP Geometric Dilution Of Precision

PWE Parabolic Wave Equation model

FD-TD Finite Difference - Time Domain

PO Physical Optics

GO Geometric Optics

DSP Digital Signal Processing

WBMOD Wide Band MODel

PM Phase Modulation

FM Frequency Modulation

AM Amplitude Modulation

LPF Low-Pass Filter

PSK Phase Shift Keying

FSK Frequency Shift Keying

MFSK Multiple Frequency Shift Keying

QAM Quadrature Amplitude Modulation

ASK Amplitude Shift Keying

AGC Automatic Gain Controller

ADC Analogue to Digital Converter

GPU Graphical Processing Unit

WAAS Wide Area Augmentation System

EGNOS European Geostationary Navigation Overlay Service

MSAS Multi-functional Satellite Augmentation System

VOR VHF Omnidirectional Radio Range

NDB Non-Directional Beacon

DME Distance Measuring Equipment

LHCP Left Hand Circularly Polarised

RHCP Right Hand Circularly Polarised

GEE British Grid Navigation System

ASC All Sky Camera

FMI Finnish Meteorological Institute

UT Universal Time

UHF Ultra High Frequency (300 MHz-3 GHz)

VHF Very High Frequency (30 MHz-300 MHz)

HF High Frequency (3 MHz-30 MHz)

MF Medium Frequency (300 kHz-3 MHz)

LF Low frequency (30 kHz-300 kHz)

VLF Very Low Frequency (3 kHz-30 kHz)

SNR Signal-to-Noise Ratio

CPU Central Processing Unit

DSB Double Side-band

SSB-SC Single Side-band Suppressed Carrier

C/N$_0$ Carrier to Noise density

ICP Integrated Carrier Phase

DFT Discrete Fourier Transform

MAT Multi-path Analysis Transform

CNES Centre National dEtudes Spatiales

HAARP High Frequency Active Auroral Research Program

Chapter 1

Introduction

1.1 Thesis Overview

This thesis describes the work undertaken to investigate the effects of the local environment and the ionosphere on a Global Navigation Satellite System (GNSS). The project focuses on the only currently fully operational system, the GPS, however consideration for the implementation of future systems has been made during development. The aims of the project are:

- To develop a simulation of the GPS which can reproduce the effects of the local environment and the ionosphere on the signal.
- To investigate the effects that the local environment and the ionosphere have on the GPS signal.
- To investigate what effect these have on the ability of the receiver to determine a position.
- To determine criteria to distinguish between local environment and ionospheric errors from the receiver outputs.

This chapter gives an overview of the thesis and the research areas of the project. The research background begins with navigation, giving a brief account of the technological advances through history and a description of the systems in use today. This is followed by an explanation of electromagnetic propagation and the effects of the environment on it. The chapter concludes with a description of the issues involved in using electromagnetic fields for communication.

Chapter 2 describes the GPS and how it provides position estimates. An overview of the whole system is given followed by more detailed descriptions of each area. A

description of the signal structure and the standard and precise (unavailable for civilian use) positioning service is given. A section describing the principle of the GPS receiver and typical methods used for positioning. The chapter concludes with a description of the causes of error in the GPS and a brief summary of the current and future methods to improve the positioning accuracy.

Chapter 3 describes the different technologies developed to improve the accuracy of GPS receivers. The start of the chapter describes the carrier-phase based surveying methods for L1 and L2 frequencies as well as methods for smoothing pseudorange measurements. The mid-section describes multi-path mitigation methods, which attempt to remove or reduce the effect of duplicate signals. This is followed by a summary of ionospheric delay correction and scintillation filtering methods. The chapter concludes with a description of the various methods that can be used to acquire the GPS signals before tracking them.

Chapter 4 describes the simulation consisting of the reference models used and the parts developed. A brief overview of a number of electromagnetic propagation models are described and assessed followed by the reasons for the method chosen. Descriptions of the IRI and WBMOD models and how they were integrated into the simulation are given. The chapter concludes with a description of how the simulator algorithm works and what is included in it.

Chapter 5 presents the results from a multi-path experiment, where a real receiver was placed in a multi-path environment and the simulator was used to reproduce the same scenario. An introduction describing the environment and how the experiment was set up, followed by the results and further investigations. A conclusion of how the simulator performed and what the results mean is given.

Chapter 6 presents further investigations into multi-path and the effects on the receiver positioning. A number of experiments were undertaken using the simulator and the results analysed for multi-path characteristics. From these results, a technique was developed to identify multi-path and it was applied to simulated and real receiver data. The chapter describes the experiments and concludes with an evaluation of the analysis technique and an explanation of the results.

Chapter 7 describes a case study undertaken following the discovery of an unusual output from a group of real receivers based in Tromsø, Norway. Three receivers showed a short but severe fade in their power output. The environment was examined to determine whether the cause was due to the local environment (multi-path) or the ionosphere. The techniques from the previous chapters were used along with outputs from other atmosphere-measuring equipment in the area to determine the cause. The simulator was used to model the effect of the sudden fade on different receiver types.

Chapter 8 provides the overall conclusions to the work and discusses what further work could be done.

Appendix A shows the user interface developed for the simulation and describes how to use it. Appendix B shows simulator outputs for different combinations of phase and amplitude scintillation.

1.2 Navigation

1.2.1 Early Navigation Methods

For thousands of years humans have found ways to navigate over the Earth. The earliest methods of navigation are thought to have been based around laying trails that could be followed. In early maritime history sailors would follow the coast and identify landmarks so as not to get lost. The first documented method of global navigation was invented in the eighteenth century and was based on measuring the angle between a charted star or the sun and the horizon. This provided good latitude accuracy but an accurate measurement of time was needed to determine the longitude due to the Earth's rotation. This problem was solved with the invention of the chronometer, the first accurate and portable clock, beginning the long standing relationship between timing and navigation.

1.2.2 Radio Navigation

In more recent history a number of systems using radio waves were implemented to calculate position. The initial systems used a ground-based transmitter and a directional antenna at the receiver to find the bearing of the transmitter. During the Second World-War a number of developments in radio navigation occurred, spurred by the need for flying at night and in bad weather; the beginnings of instrument flying. The first long-range navigation system was the British Grid Navigation System (GEE short for grid) developed and used by the British air force. This system could provide the position relative to the transmitter stations rather than just a bearing. GEE used the Very High Frequency (30 MHz-300 MHz) (VHF) band which limited it to line of sight propagation; this produced accurate measurements but was vulnerable to being blocked. The United States developed a system based on the principles of GEE known as the LOng RAnge Navigation system (LORAN). It used the Medium Frequency (300 kHz-3 MHz) (MF) band giving it greater range and less vulnerability to obstacles, but it was less accurate. The principle behind the systems was to find the time difference between the arriving signals from a number of transmitters of known location. The transmissions were synchronised so it was possible, based on the difference in time between two

signals, to plot a hyperbolic curve of equal delay difference between the two transmitters on a map (see Figure 1.1). With three transmitters a location could be determined from where the curves crossed on the map.

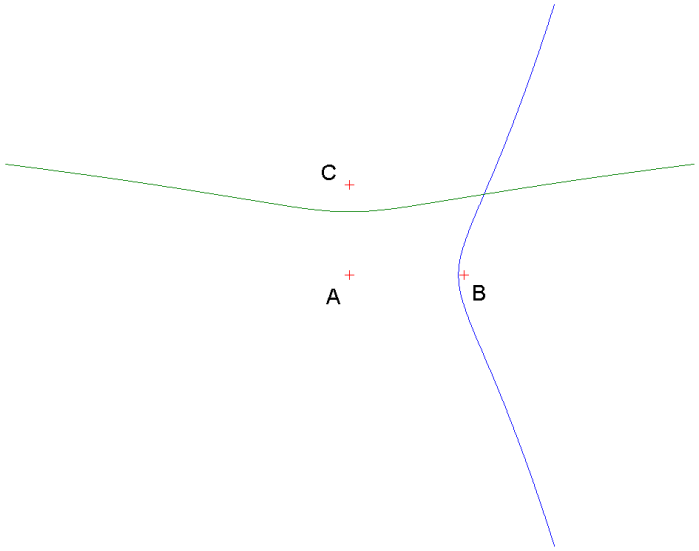


Figure 1.1: Diagram showing the positioning method of GEE. The curves have an equal delay difference between two stations: the green curve is between station A and C; the blue curve is between station A and B. Where the curves intersect determine the position

The Omega navigation system, developed by the United States, was the first truly global navigation system. It used the Very Low Frequency (3 kHz-30 kHz) (VLF) band, which made efficient long distance propagation possible. This gave the advantage of only requiring eight antenna stations to provide global coverage. Unfortunately, the disadvantage of using very long propagation distances is that the signals can be significantly delayed by the atmosphere causing position errors of 2 to 4 nautical miles. The system worked in two modes: direct ranging and hyperbolic. The direct ranging mode was a differential system that could only provide the position change from an initial point. This was done by measuring the phase change between the received signal and its own reference oscillator, where one wavelength was approximately 25 km . Using this method only two stations were required to track the position, however the reference oscillator was critical to the operation and failure or semi-failure could be catastrophic. The hyperbolic mode measured the difference between two signals from different transmitters providing hyperbolic lines of equal delay difference, similar to GEE and LORAN. This method could provide a position fix within the wavelength distance (25 km), however there was an ambiguity of the number of additional wave-

lengths offset. This meant that if a reasonable idea of the location was known the system could be initialised. Three transmitters were required to obtain a position using this method, however there was no reliance on a reference oscillator.

1.2.3 Satellite Navigation

To overcome the problems associated with ground-based systems the idea of using artificial satellites to transmit the signals required for positioning was proposed. The first operational system was called Navy Navigation Satellite System (NAVSAT) or "Transit". It consisted of five satellites (Figure 1.2) in low level polar orbits to minimise the orbit period. The satellites broadcast a signal containing the precise time and orbit parameters. The receiver position was determined by the Doppler measurements taken as the satellite passed. The gradient of the Doppler measurements determined the perpendicular distance (longitude) from the satellite track where as the point at which the Doppler crossed zero determined the position along the satellite track (latitude). The advantage of the system was availability due to the use of satellites. The disadvantages were the sensitivity to receiver movement and the time required for a position fix.

A number of other systems were developed with some success but the next evolution in global navigation came with the "Navstar" GPS and the Globalnaya Navigatsionnay Sputnikovaya Sistema (GLONASS). GPS was developed by the United States Department of Defence (DoD) and became fully operational in 1994. The GLONASS was developed by the former Soviet Union and became fully operational in 1995 but is now managed by the Russian Federation Ministry of Defence. The GPS is the only fully operational system at the time of writing. These systems have a number of satellites (GPS satellite in Figure 1.3) in orbit above the Earth, each transmitting a different signal. The GPS signal comprises of a code that identifies the satellite and information about the satellite position, corrections and time of transmission. The receiver combines the information with techniques to improve the timing accuracy and provides a position fix using triangulation. Further details of the GPS are provided in the next chapter.

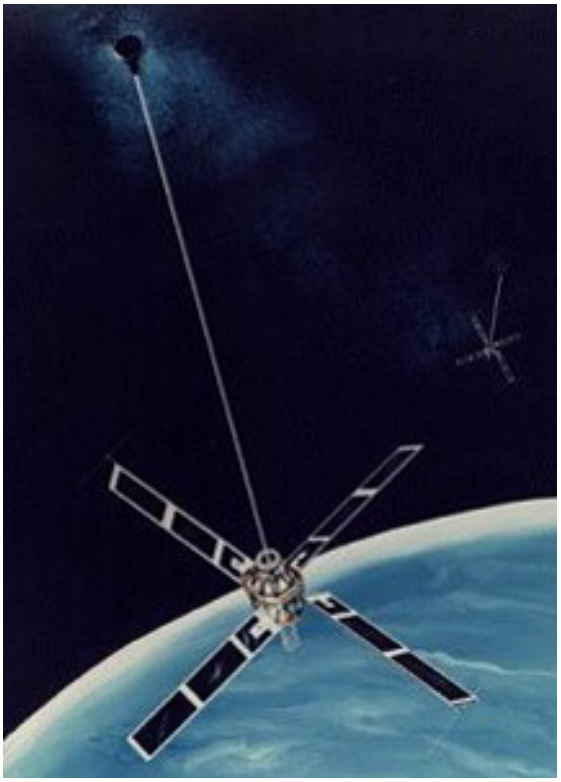


Figure 1.2: Image of “Transit” satellite courtesy of US Navy

Figure 1.3: Image of GPS satellite courtesy of NASA

1.2.4 Current Radio Navigation Systems

Although GPS has now become the main navigation aid, for some applications there are still occasions when ground-based systems are preferable. One example is in aviation where the required accuracy and integrity is more than the GPS Standard Positioning Service (SPS) can currently provide. There is an upgrade available for use in the United States called the Wide Area Augmentation System (WAAS) that provides ionospheric corrections to receivers to increase accuracy (described in Chapter 2). There are also European and Asian upgrades under way called European Geostationary Navigation Overlay Service (EGNOS) and Multi-functional Satellite Augmentation System (MSAS) respectively. These provided the same type of corrections as the US WAAS system, but for the different regions. Until the correction systems are in wide use, navigation aids using ground-based technology are relied upon.

The two main aircraft navigation systems in use today are the VHF Omnidirectional Radio Range (VOR) and Non-Directional Beacon (NDB) systems: The NDB system is the older and less precise of the two; it is based on a number of omni-directional beacons on the ground and a narrow directional antenna on the aircraft. The beacons are modulated with a two letter Morse code so that each one can be identified on an aircraft radio. The directional antenna on board the aircraft spins in azimuth giving a stronger output when the antenna is pointing toward a beacon. This provides a bearing to the beacon but is dependent on the aircraft orientation, so compensating for wind or obtaining a position fix is difficult.

The VOR system is also based on a series of beacons, however it can provide a bearing which is independent of the aircraft orientation using two signals. The first signal is transmitted omni-directionally and is modulated with a Morse code identifier and a 30 Hz reference oscillator. The second signal consists of a carrier transmitted by a rotating directional antenna that revolves at 30 Hz . This produces two 30 Hz carriers that have a different phase for each bearing from the receiver. The receiver derives the phase difference between the two 30 Hz carriers to obtain the bearing (known as the radial) from the beacon. Often combined with the VOR beacon is Distance Measuring Equipment (DME) where the range from the beacon is determined by a round trip propagation of a signal from the aircraft to the beacon.

1.3 Radio Propagation

1.3.1 Electromagnetic Radiation

The discovery of electromagnetic radiation can be attributed to a number of people. Although Maxwell formed the set of equations that describe electromagnetism, his work

was based on a number of peoples previous achievements:

Coulomb's law

Coulomb discovered that the electric field emanating from an object is directly proportional to the charge contained within it and inversely proportional to the square of the distance from the object.

$$\hat{E} = \frac{Q_1}{4\pi\epsilon r^2}\hat{r} \tag{1.3.1}$$

where,

- $\hat{E}$ is the electric field strength (Vm^{-1}).
- Q_1 is the electric charge (C).
- ϵ is the permittivity (Fm^{-1}).
- r is the distance from the charge (m).
- $\hat{r}$ is the unit vector representing the direction of the electric field.

Ampere's law

Ampere's experiment consisted of a current carrying wire and a compass. He noted that when the compass was brought near the wire , the needle followed a circular path around the wire. Ampere derived that the current through a conducting wire produces a constant magnetic field flowing in closed loops on a plane perpendicular to the wire.

$$I = \oint H.dL = \int_s J.dS \tag{1.3.2}$$

where,

- I is the electric current (A).
- H is the magnetic field (Am^{-1}).
- L is the line length integrated over (m).
- J is the current density (Am^{-2}).
- S is the surface area integrated over (m^2).

Gauss' Law

Gauss' law is the specific application of Gauss' Theorem to electrostatics and magnetostatics. The electrostatic case states that the integral of the electric flux over a closed surface is equal to the enclosed charge divided by the permittivity. Therefore whatever the size of the volume enclosed by the surface the total flux flowing out is equal. However the electric field strength is inversely proportional to the area of the enclosed surface, so for a sphere it is proportional to $\frac{1}{4\pi r^2}$ (Coulomb's law).

$$\Phi = \oint_s E.dS = \frac{Q}{\epsilon} \tag{1.3.3}$$

where,

- Φ is the electric flux (Nm^2/C).
- E is the electric field strength (Vm^{-1}).
- S is the surface area integrated over (m^2).
- Q is the electrical charge (C).
- ϵ is the permittivity (Fm^{-1}).

The magneto static case states that the magnetic flux flowing through a closed surface is always zero. This is proved as magnetic fields always form closed loops, there is no magnetic monopole (or magnetic charge) currently known of.

$$\oint_s B.\delta S = 0 \tag{1.3.4}$$

Faraday's law

Ampere showed that a constant electric current produces a magnetic field, however Faraday found that a constant magnetic field does not produce a current. In order to produce a current the magnetic field must be continually changing. Faraday discovered that by moving a magnet near a wire coil, a current was induced. Faraday's law states that the electromotive force generated in a stationary wire loop equals the surface integral of the time rate of change of the magnetic flux density integrated over the loop area.

$$V = \oint E.dL = -\int\int \frac{\delta B}{\delta t}.dS \tag{1.3.5}$$

where,

- V is the voltage (V).

- E is the electric field strength (Vm^{-1}).
- L is the line length integrated over (m).
- B is the magnetic flux density (T).
- t is the time (s).
- S is the surface area integrated over (m^2).

Maxwell's Equations

Maxwell deduced from the above laws that electric and magnetic fields were not independent and that a dynamic field combining the two static fields was possible. The dynamic field is in the form of a wave, known as an electromagnetic wave. The advantage that electromagnetic waves provide over static electric or magnetic fields are that they can travel much further. Static electric fields degrade proportionally to the square of the distance travelled, where as electromagnetic waves degrade proportionally to the distance travelled. The oscillation of electromagnetic wave causes a perpetual effect, where the electric field generates the magnetic field, which in turn generates the electric field enabling the longer propagation. Maxwell combined the previous work with his own to form what are known as the Maxwell field equations which describe electromagnetic propagation.

$$\nabla.E = \frac{\rho}{\epsilon_0} \tag{1.3.6}$$

$$\nabla.B = 0 \tag{1.3.7}$$

$$\nabla \times E = -\frac{\delta B}{\delta t} \tag{1.3.8}$$

$$\nabla \times B = \mu_0 J + \mu_0 \epsilon_0 \frac{\delta E}{\delta t} \tag{1.3.9}$$

where,

- B is the magnetic field strength (T).
- E is the electric field strength (Vm^{-1}).
- J is the current density (Am^{-2}).
- t is the time (s).
- μ_0 is the magnetic permeability of the medium (NA^2).
- ϵ_0 is the electric permittivity of the medium (Fm^{-1}).
- ρ is the volume charge density (Cm^{-3}).

The first two equations are based on Coulomb and Gauss law for electric and magnetic fields. These give the static relations where the electric and magnetic fields can be considered separately, which can also hold true for slow changing fields. The third and fourth equations are for dynamic fields and are based on Faraday's law and a part of Ampere's law. From these equations, Maxwell deduced that electromagnetic waves travel at the speed of light and this combined with the properties of light showed that light is an electromagnetic wave. This provided a fundamental understanding of electromagnetic waves and of what light is, which led to the beginning of quantum physics and Einstein's theory of relativity. The equations suggested that other non visible waves exist, which was proven correct with the discovery of other parts of the electromagnetic spectrum like radio waves, microwaves, x-rays and gamma rays. Heinrich Rudolf Hertz was the first to prove the existence of radio waves, building circuits that transmitted and received them. This led to Marconi's inventions which used the radio waves to carry information, making radio communication possible. The inventions of radio, television, radar and satellite communication all have their origins in Maxwell's electromagnetic theory.

1.3.2 The Atmosphere

The atmosphere is a gaseous layer surrounding the surface of the Earth. The atmosphere can be sub-divided into layers according to temperature profiles, see Figure 1.4. However, in the field of electromagnetic propagation another property is important, the electron density. Free electrons are formed in the region of the atmosphere that is susceptible to the ionising radiation from the Sun. The region is known as the ionosphere and it covers an altitude of 50 km to 600 km . The two most important layers for electromagnetic propagation are the ionosphere due to its charged state, and the troposphere due to its changes in density and weather.

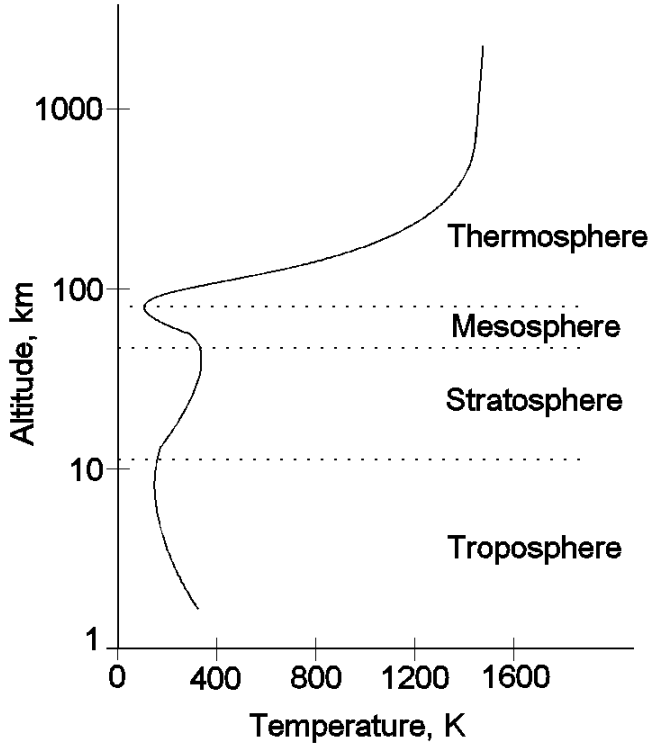


Figure 1.4: The Thermal Layers of the atmosphere

The Troposphere

The troposphere is the lowest region of the atmosphere and contains a dense mixture of gases. It contains mainly nitrogen and oxygen with a small percentage of other gases. The troposphere begins at the Earth's surface and ends at approximately 12 km above the earth. All meteorological weather as well as particles such as dust and ash are contained within the troposphere. The temperature decreases with height until the tropopause which is a thin transition layer to the stratosphere, where the temperature begins to increase. The troposphere effects electromagnetic waves to varying degrees depending on the frequency of the wave. At microwave frequencies, the most significant effect is rain and atmospheric fading. Rain fading causes electromagnetic waves to weaken and scatter due to refraction, diffraction and absorption from the rain drops. The severity of the effect is determined by the rain density and the frequency of the microwave. Rain fading is completely dependent on weather and is difficult to predict in most circumstances. Atmospheric fading is determined by the the properties of the molecules in the atmosphere and is less variable than rain fading. The effects of atmospheric fading at microwave frequencies are absorption and refraction mainly from oxygen and water vapour. The electromagnetic absorption of oxygen and water vapour

over the microwave spectrum is shown in Figure 1.5. Both rain and atmospheric fading cause changes in refractive index which can add delays to propagation time.

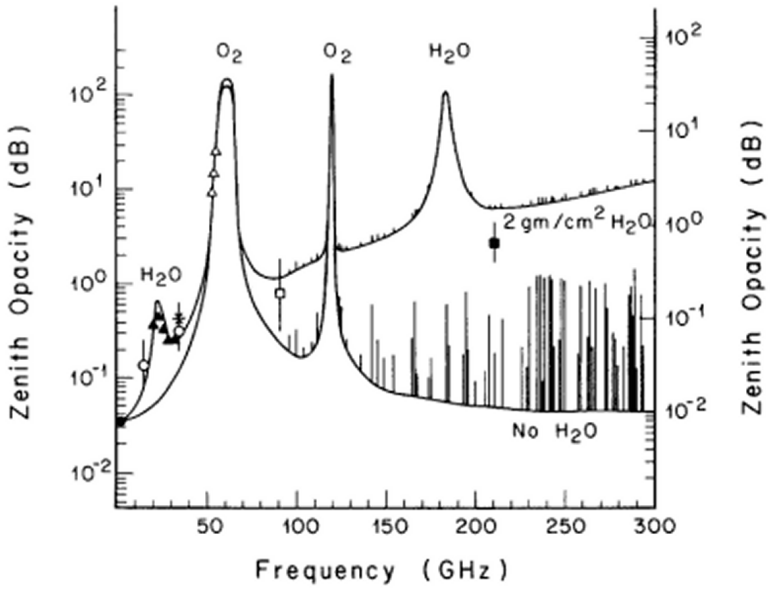


Figure 1.5: Figure showing the attenuation of electromagnetic radiation by standard atmosphere and water vapour (Copyright 2006 - CNES).

The troposphere has a significant refractive effect on VHF frequencies causing them to travel further than line of sight. The waves refract back towards the earth as the refractive index of the troposphere reduces with height (it is directly proportional to the pressure). The refractive index is also inversely proportional to temperature so the refraction effect can be significantly increased during a temperature inversion. A temperature inversion is a weather phenomena where the temperature increases (or remains constant) as the altitude increases, rather than the expected decrease. This causes a duct of relatively constant refractive index with sharper changes above and below the inversion. This duct channels the electromagnetic waves with the curvature of the earth allowing a much greater propagation distance.

HF frequencies can be subject to sky-wave propagation, where signals refract between the ground and the ionosphere. This is possible because of the three significantly different refractive indexes of the; Earth, troposphere, ionosphere. This enables longer distance propagation and is often used in communication channels. LF frequencies can travel great distances by ground wave propagation where it follows the curvature of the earth by diffracting over it. The wavelength of LF frequencies in comparison to the curvature of the earth allows this. The conductivity of the ground determines the strength of the effect. As mentioned in the Navigation section, Omega was able to provide global coverage with eight LF antennas.

The Ionosphere

The ionosphere layer covers an altitude of 50 km to 600 km . It is composed of oxygen and nitrogen molecules at lower altitude regions and oxygen, helium and hydrogen atoms at higher regions. Radiation from the Sun's rays cause the electrons in the atoms to move to higher energy states until they break away completely (ionise). This causes the gas to become a plasma where ions and electrons are continually formed and recombined back to a neutral state. At the higher altitude regions many are ionised and few recombine, however the molecule density is low therefore the ion density is also low. The lower regions are less exposed to the ionising radiation and the molecule density is high therefore few are ionised and most recombine rapidly. The mid regions are the most significant to radio propagation as the molecule density is high enough for substantial ionisation, but low enough to increase recombination time giving a permanent plasma state. This balance of ionisation to recombination varies with altitude and time, the typical shape of the ion density can be seen in Figure 1.6.

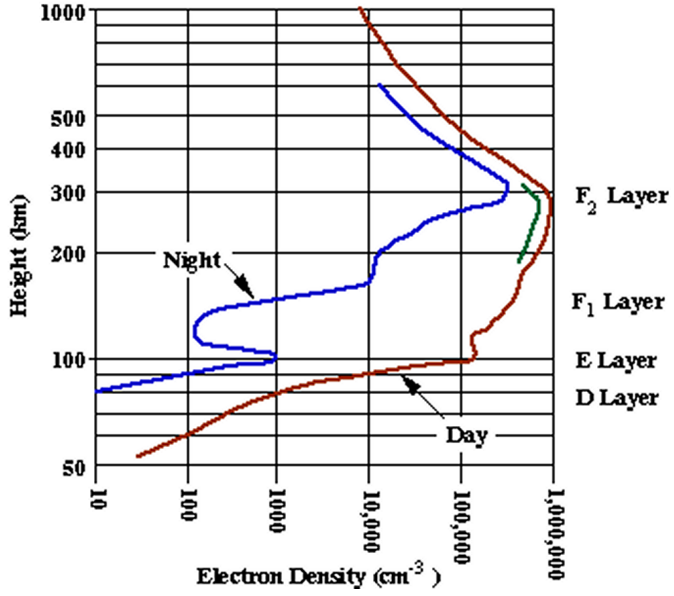


Figure 1.6: Figure showing the electron density verses altitude in the ionosphere, courtesy of HAARP

The ionosphere can be sub-categorised into further layers, shown in Figure 1.6, known as the D, E, F_1, F_2 regions :

- The D region is the lowest region beginning at an altitude of 50 km and ending at 90 km . The rate of ionisation in this region is low as many of the Sun's photons (of the correct energy) have already been absorbed by higher regions. This region can refract LF signals back to earth, however higher frequencies pass through it. The D region only exists during the day because of the low amount of ionisation

and its high molecule density where the ions can be rapidly recombined out of the Sun's rays.

- The E region covers an altitude of 90 km to 140 km . The ionising rate in this layer is greater than the D region enabling the refraction of higher frequency radio waves. The E region can refract signals with frequencies of up to 20 MHz back to earth during the day. At night the E region can maintain a plasma for a longer period of time than the D region as the molecule density is lower (reducing recombination) and there are more ions. Typically, all the ions that form the plasma recombine by midnight.

- The F_1 and F_2 regions cover an altitude of 140 km to 210 km and 210 km to 600 km respectively. They exist as separate regions during the day but combine into one region at night. The ionisation rate is quite high in these regions and the molecule density is lower than D and E regions causing a sustained plasma through day and night. The F regions can refract radio waves up to 30 MHz (HF) back to earth, commonly used for sky-wave propagation mentioned earlier. Although radio frequencies greater that 30 MHz can pass through the ionosphere, they are still refracted.
(*Miller*, 1996)

As mentioned the ionosphere can refract radio waves and its refractive index is proportional to the electron density. When a radio wave propagates up through the ionosphere it is refracted towards earth, until it reaches the F_2 peak (see Figure 1.6), where the refractive index begins to reduce causing the wave to refract away from the earth. As mentioned in the previous section at certain frequencies the radio wave cannot reach the F_2 layer and is refracted all the way back to earth (sky-wave propagation). The change in refractive index also causes a change in the radio wave velocity introducing a delay in the propagation time. The ionosphere does not have a smooth change in refractive index (however it can be assumed for refraction and delay effects) but is full of small fluctuations due to the varying rates of ionisation and recombination. These small scale changes cause parts of the electromagnetic wave to scatter which cause rapid changes in wave interference. This interference appears as fluctuations in the phase and intensity of the wave at a receiver and is referred to as scintillation.

1.3.3 Electromagnetic Wave Interactions

Radio propagation is not only affected by the gaseous atmosphere but also by liquid and solid structures on the surface of the earth. The chemical composition of the structure determines how much it refracts and/or reflects the signal where as the shape of the structure is the main property that determines the diffraction characteristics. If a

signal passes through an aperture or an edge of the same order size as the signals wavelength it causes it to spread around the structure. From Maxwell's equations, the three properties that influence the behaviour of the wave interactions are the conductivity, permittivity and permeability of the material. These properties determine whether the material acts as a conductor, where nearly all the wave is reflected, or an insulator, where some of the wave is reflected and some is refracted. All of these effects cause the radio waves to travel different paths to the destination (a receiver in most cases) and is known as multi-path, see Figure 1.7(i). The convergence of the signals at the receiver can cause interference, where the received signal is distorted from the signal transmitted. The interference can affect the carrier and modulation differently. The effect on the carrier can be constructive or destructive depending on whether the majority of the carriers are in-phase (equal to a multiple of the period) or out of phase (half a wavelength difference from in phase), see Figure 1.7(ii). Multi-path also effects the signal modulated on to the carrier, where the multiple copies can cause distortion or other effects depending on the type of modulation. Digital modulation types have been recognised as more robust to multi-path than their analogue counterparts as the modulation is at distinct points and discrete values therefore signal processing techniques can be used to remove the repeated signals. This is largely responsible for the current digital revolution in radio and television broadcasting. An example of the multi-path problems associated with analogue modulation techniques are the double images (known as ghosting) that can be seen on analogue televisions in built up areas.

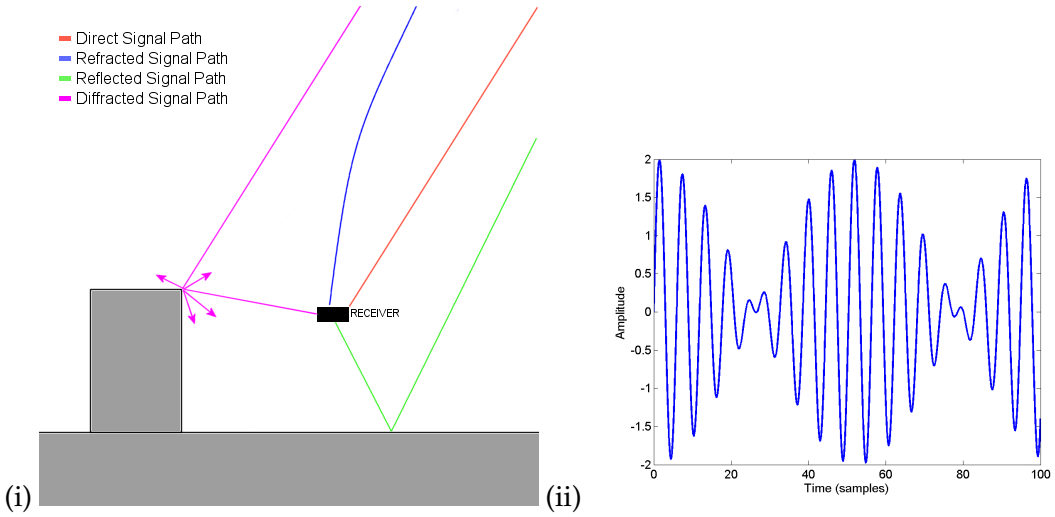


Figure 1.7: (i) Diagram showing the possible multi-path routes from transmitter to receiver; (ii) Plot showing the effect of constructive and destructive carrier interference

1.4 Radio Communication

1.4.1 Analogue Modulation Techniques

As discussed in the electromagnetic propagation section, electric and magnetic fields only propagate efficiently when combined as an oscillating wave. Therefore, in order to transmit an irregular signal, it must be combined with an oscillating signal. The oscillating signal is known as the carrier and the signal can be modulated on to the carrier using one of a number of techniques.

Amplitude Modulation

In the 1920's the invention of the vacuum tube led to the first transmission of sound over a radio link using Amplitude Modulation (AM). It is the most basic technique, where the strength of the transmitted carrier is varied corresponding to the signal to be sent. This is done by mixing (or multiplying) the two signals together (see Figure 1.8). In the frequency domain this causes a single carrier and an upper and lower side-band, seen in the third graph of Figure 1.8.
There are many methods to demodulate the signal back into its original form, the most simple is based on a diode detector. A bandpass filter is used to select the frequency band, the diode removes the negative parts of the waveform and a low pass filter removes the carrier. As the diode gives a fairly linear response on the positive parts of the waveform a reasonable quality is obtained. The super heterodyne method provides better quality as it uses an oscillator to select frequencies enabling the use of better constant frequency filtering and amplifying components.
One of the main issues with AM is that any loss in carrier amplitude during propagation directly effects the recovered signal. This loss in amplitude can be due to distance, obstacles, atmosphere or multi-path. This is responsible for the changes in volume that can occur when listening to AM radio stations. The effect of distance and slow changes in the received signal strength can be compensated for by an Automatic Gain Controller (AGC) in the receiver. This attempts to maintain the signal strength by varying the amplification according to the peak amplitude of the received signal.
In AM each side-band contains all the information required to reconstruct the original signal. Therefore, it is possible to reduce the required bandwidth and transmit power by removing one side-band and the carrier by filtering. This is a technique is known as Single Side-band Suppressed Carrier (SSB-SC) and it halves the required bandwidth and quarters the required transmit power.

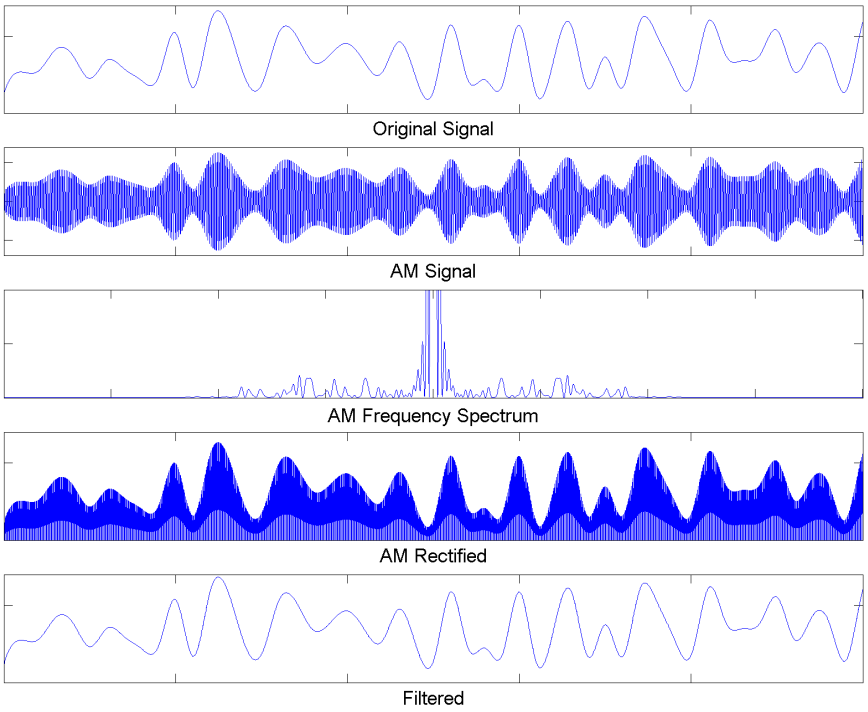


Figure 1.8: Figure showing an amplitude modulated signal and its frequency spectrum

Frequency Modulation

Frequency Modulation (FM) and Phase Modulation (PM) were developed as a way to remove the relation between the modulated signal amplitude and the carrier amplitude, due to the unpredictable carrier losses during propagation. Rather than vary the amplitude of the carrier in proportion to the signal, it varies its frequency or phase as shown in Figure 1.9 and 1.10 respectively. A common way of doing this is by using a Voltage Controlled Oscillator (VCO) which produces an oscillating wave with a frequency proportional to an input voltage.

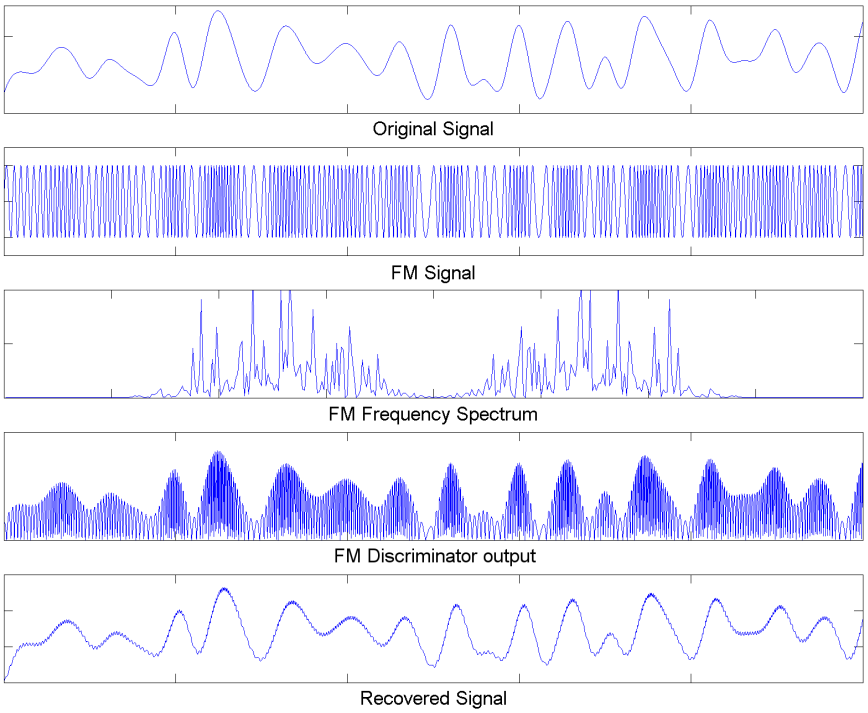


Figure 1.9: Figure showing a frequency modulated signal and its frequency spectrum

The signal can then be mixed to any frequency, determined by propagation characteristics or licence. As mentioned, the mixing process creates duplicate images of the modulated signal so these must be filtered before transmission. The amplitude of the carrier remains constant at the transmitter, however it can be affected in the same way as AM during propagation. Similarly to AM, the FM receiver can use an AGC to maintain a constant amplitude. However the AGC has a constant amplitude reference to monitor where as the AM AGC must monitor the peak outputs. To recover the original signal, the received signal is mixed back to the lower frequency. At the lower frequency, a discriminator can be used to convert it back to the original signal. The PM discriminator converts a phase change in the oscillating wave into an amplitude change. A typical discriminator is based on a Phase Locked Loop (PLL) which is a closed loop system that tracks the changes in the frequency or phase and outputs the changes. A full description of the PLL is provided in the next chapter.

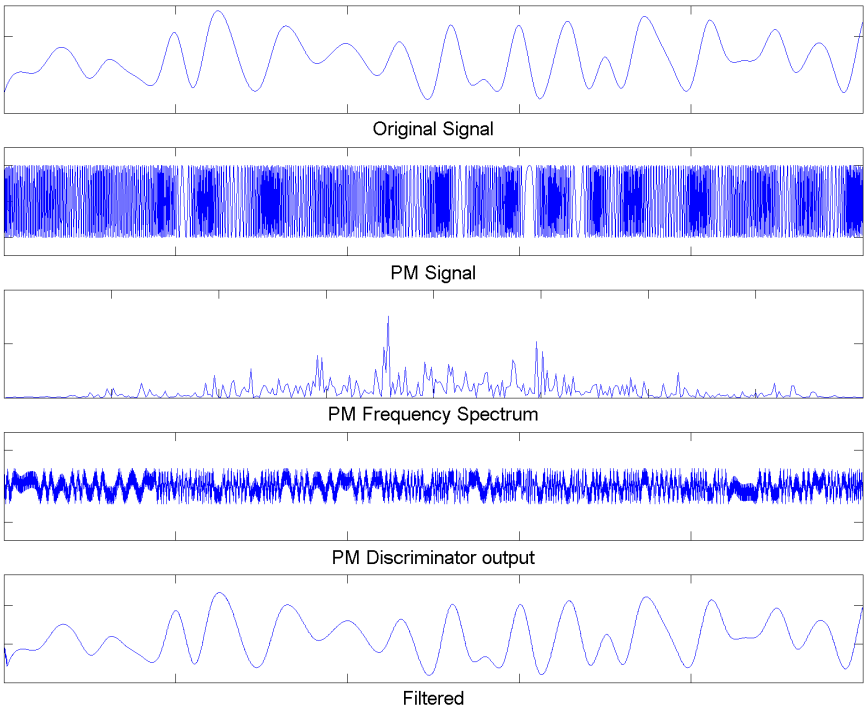


Figure 1.10: Figure showing an phase modulated signal and its frequency spectrum

1.4.2 Digital Modulation Techniques

Digital Modulation techniques rely on similar methods to their analogue equivalents; amplitude, frequency and phase modulation.

Amplitude Modulation

This type of modulation is based around Amplitude Shift Keying (ASK) where two or more levels of amplitude represent digital values, shown in Figure 1.11. The simplest two level method is implemented by turning the transmitter on and off. As the number of levels increase, the likelihood of an error increases, which is especially true for ASK. When using more than two levels each level is known as a symbol, as the baud rate is no longer the same as the rate of level changes. So if four symbols (levels) are used, then one symbol represents the same as two bits (one of four values) making the baud rate twice that of the symbol rate. The transmission stage is similar to the analogue method but the receiver stage is much more critical when using this method. The receiver must have an AGC to amplify the signal to a constant so that each level corresponds to the correct value. A typical AGC works by maintaining the peak output constant, so if a

peak value is not received for some time a lower peak is used shifting the levels and causing errors. These issues deter the use of multiple level ASK in most applications. To demodulate, the receiver rectifies and filters the signal then detects the level using thresholds. A method called Quadrature Amplitude Modulation (QAM) can extend the number of levels used by modulating two carriers 90° out of phase. The two level QAM method produces a waveform which switches between the two carriers, producing a similar signal as the Binary Phase Shift Keying (BPSK) method described later.

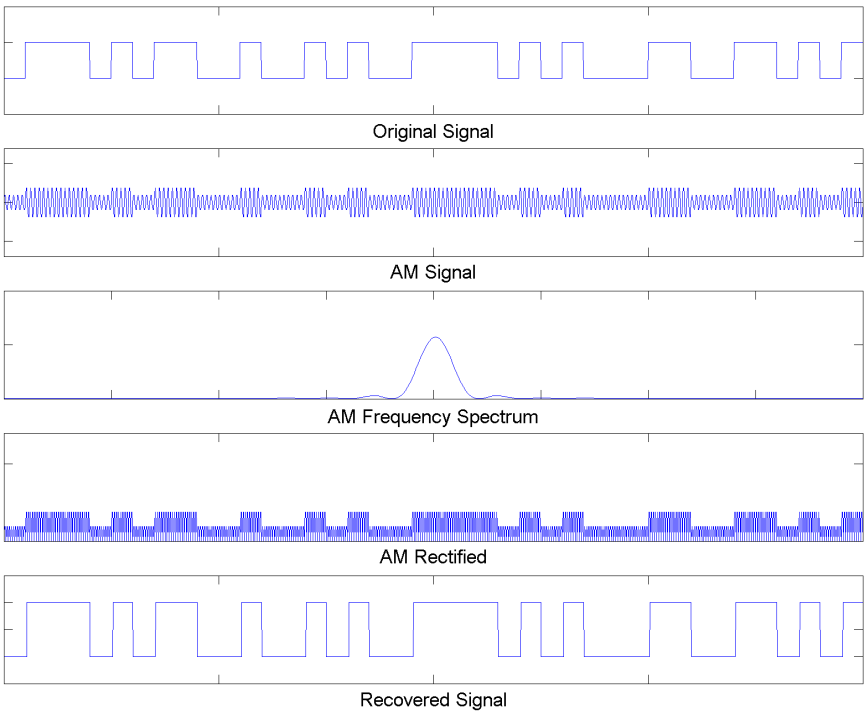


Figure 1.11: Figure showing Amplitude Shift Keying Modulation

Frequency Modulation

The basic form of this modulation is Frequency Shift Keying (FSK) where two frequencies are used to represent a one and a zero of a digital message, shown in Figure 1.12. The typically separation between each frequency used is equal to the bandwidth of the data to be modulated. There is a non-coherent method where the frequency change is performed instantaneously causing abrupt phase changes during switching. The coherent method relates the frequency separation of the modulation frequencies to the data rate, where the changes are performed when the two frequencies are at equal phases. This gives a continuous phase waveform output which reduces the er-

ror rate as sharp phase shifts cause wide bandwidth noise. Multiple Frequency Shift Keying (MFSK) is the same method but uses more than two frequencies. Demodulation is achieved by a bank of filters, each centred on one of the frequencies used, then the outputs are integrated providing a peak at the correct frequency. There are also methods based in DSP that use the FFT to both modulate and demodulate the signal.

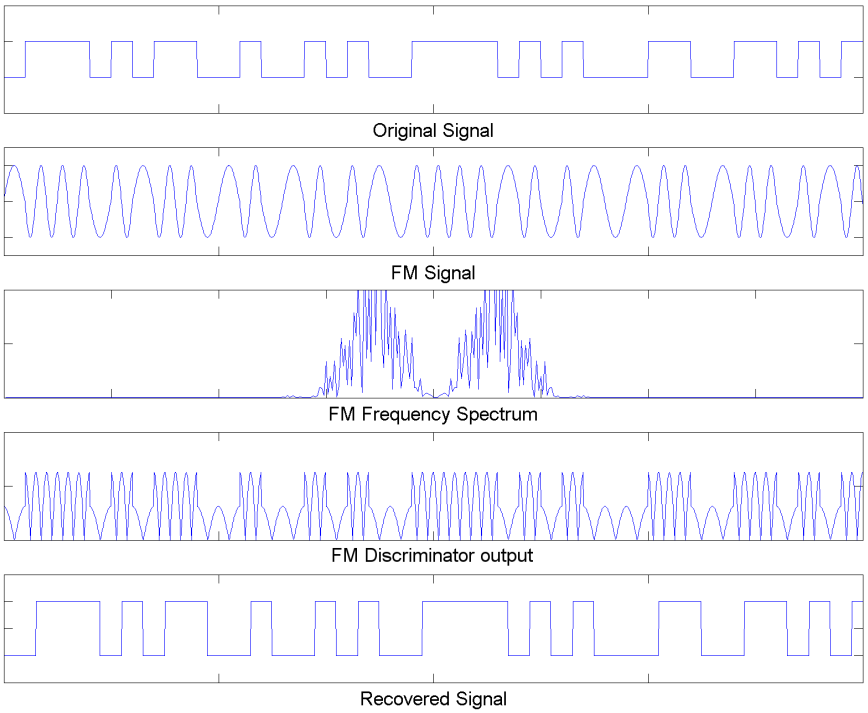


Figure 1.12: Figure showing Frequency Shift Keying Modulation

Phase Modulation

Phase modulation is based on the Phase Shift Keying (PSK) method where the phase of the carrier is shifted to represent different digital values, shown in Figure 1.13. The simplest of these is BPSK which uses two phases of the carrier, 180 ° apart, to represent the two values of the digital signal. Quadrature Phase Shift Keying (QPSK) uses four phases separated by 90 ° to encode the data, enabling two bits of data to be sent per symbol. PSK has been used with more than 16 symbols, but this increases the probability of an error. The BPSK modulation can be generated by switching between an inverted and a non-inverted carrier. Typically a type of PLL is used to demodulate the signal. A Costas PLL can be used for demodulation of BPSK and a modified version can demodulate QPSK. The loops work by using two PLLs one to track the 0 ° offset

carrier and one for the 180 ° offset carrier. By combining the two PLL a new error slope is formed with two stable points separated by 180 ° . These methods are described fully in the next chapter.

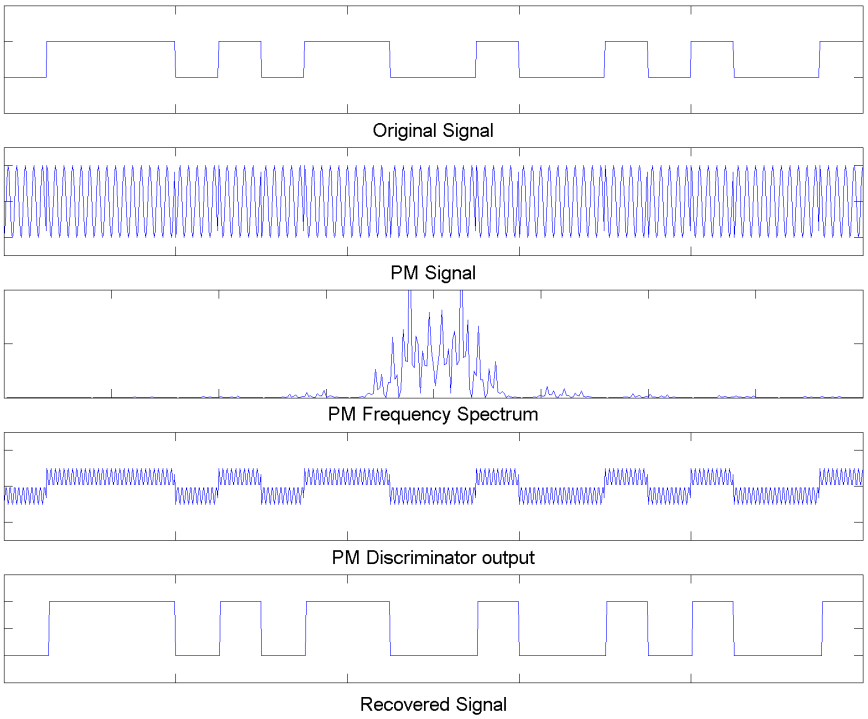


Figure 1.13: Figure showing Phase Shift Keying Modulation

1.4.3 Multiplexing and Multiple Access

There are a number of techniques that can be used to send multiple signals across a single communication channel. Multiplexing is a method of combining multiple signals into one signal transmitted from one station. Multiple access is a method of allowing multiple stations to transmit their modulated carrier using the same communication channel (frequency band and medium). The two methods accomplish this is similar ways.

Time Division Multiplexing and Multiple Access

Time Division Multiplexing (TDM) divides the different signals into time segments and transmits the segments from each signal in turn. The receiving station separates the signal back into segments and reforms the original signals. Time Division Multiple

Access (TDMA) is similar, but each station is assigned a time slot that they can broadcast during. An accurate time reference is required at each station so transmissions do not overlap.

Frequency Division Multiplexing and Multiple Access

Frequency Division Multiplexing (FDM) assigns a bandwidth to each signal and transmits each signal on a separate carrier frequency. The receiving station demodulates each frequency reforming the original signals. Frequency Division Multiple Access (FDMA) assigns a frequency and bandwidth limit to each station to ensure the signal bandwidths do not overlap.

Direct Sequence Code Division Multiplexing and Multiple Access

Direct Sequence CDM (DSCDM) uses a digital code to represent each signal, where each code has a low cross-correlation with the other codes. The signal amplitude is multiplied by the digital code, where the code repeat frequency is at least the bandwidth of the signal (effectively sampling the signal). This spreads out the signal frequency spectrum making the signal indistinguishable from background noise, however it is now possible to transmit other signals with different codes over the same frequency band. The receiver must remove the code of the required signal by matching it with a replica and then integrate for the length of the code. This correlates the two codes (known as despreading) and recovers the signal. Direct Sequence CDMA (DSCDMA) uses the same method but assigns a code to each transmitting station.

Frequency Hopping Code Division Multiplexing and Multiple Access

This method is similar to the direct-sequence method, but rather than shifting the phase of the carrier, it changes the frequency. The changes in frequency must be determined by algorithm or code so the receiver is able to predict the changes. Data is modulated on each frequency using any of the digital modulation techniques. If the multiple access method is used, the possibility of the different transmitters broadcasting on the same frequency at the same time must be prevented or handled.

1.5 Summary

This chapter has described the thesis and the aims of the project. An overview of navigation is given, showing the progression from early methods to the satellite based

systems used today. The overview shows that although the technology has advanced rapidly, the principles are still very similar. A section describing general electromagnetic propagation and how it is affected by the atmosphere and local environment is shown. Finally, the methods used to carry information on electromagnetic waves have been introduced.

Chapter 2

The Global Positioning System

2.1 Introduction

This chapter provides a description of the GPS satellite navigation system; an overview of the system is first given followed by a detailed description of the critical areas. The signal structure is described as well as how a typical receiver uses it to determine the position. The causes of errors in the GPS are described and how they affect the signal and the positioning accuracy. The chapter concludes with a discussion of the future of satellite navigation.

2.2 Overview

The current GPS system is based around thirty-one orbiting satellites (at time of writing), each transmits signals over the surface of the Earth. GPS receivers determine the time taken for the signal to reach the receiver from each satellite. This time delay is directly proportional to the separation distance and can be used to calculate the position using a technique known as triangulation. The GPS consists of three main parts-

- The space segment
- The control segment
- The user segment

2.2.1 The Space Segment

The space segment consists of satellites that broadcast the timing signals down to earth. The satellites are in orbit at an altitude of approximately 20,000 km ; one orbit takes ap-

proximately twelve hours to complete. The satellites are split into six orbital planes inclined at 55 ° to the equator shown in Figure 2.1. The satellites themselves are powered by on-board batteries charged by solar panels when in view of the sun. As the GPS is based on timing information, it is necessary for the satellite timing source to be very accurate therefore an atomic clock is used, which provides an accuracy of 1 part in 10^{13}. Each satellite contains multiple clocks for redundancy, typically four, two based on rubidium isotope and two on caesium detailed in *McCaskill et al.* (1995). The satellites are attitude stabilised in three axes to keep the antenna pointing at the centre of the earth. The antenna positioned on each satellite is a helix-element array, the helix elements polarise the signal and the array shapes the beam. The beam shape is similar to that shown in Figure 2.2, where the gain is reduced in the centre and increased around the sides. This is done to maintain a similar signal intensity over the Earth's surface.

The satellites are subject to a time dilation effect due to their velocity, predicted by Einstein's theory of general relativity. As this is a near constant effect the problem was reduced by a slight increase in the frequency of the carrier wave, which when time dilated reduces to the correct frequency. The signal is broadcast on a 1575 MHz (plus the time dilation offset) carrier and is Right Hand Circularly Polarised (RHCP), where the polarisation rotates clock-wise at a rate equal to the carrier frequency. Linear polarisation would be impractical for GPS, as it requires the receiver antenna to be aligned to the transmit antenna(e.g. Dipole antenna), which in this case can be in multiple and constantly changing positions. Circular polarisation also helps to reject multi-path as reflections can cause the polarisation to invert from RHCP to Left Hand Circularly Polarised (LHCP). Typical receiver antennas can reject LHCP waves by 3 dB to 10 dB, however if the signal is reflected twice it reverts back to a RHCP wave and no rejection occurs.

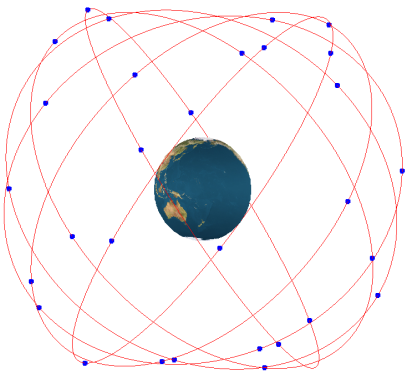


Figure 2.1: Figure showing the orbital constellation of the GPS

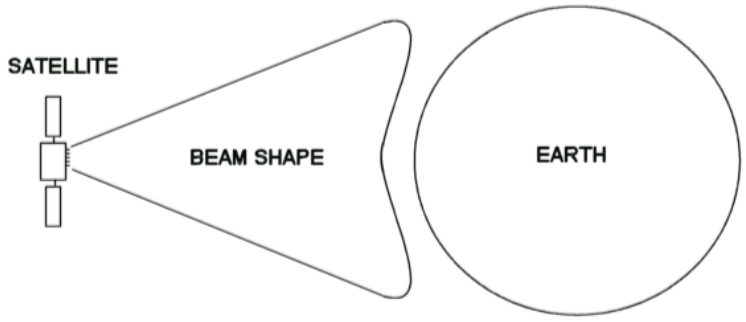


Figure 2.2: Approximation of the beam shape from a GPS satellite antenna

The GPS provides two services, the Standard Positioning Service (SPS) and the Precise Positioning Service (PPS), the SPS is unrestricted and offered for civilian use where as the PPS is restricted at the discretion of the United States DoD. The SPS uses 1023 chip length codes to access the system, it is a single-frequency system and has a horizontal accuracy of approximately 12 metres. The PPS uses 15345037 chip length codes to access the system; it uses dual-frequency system with a narrower code chip-width which improves the horizontal accuracy to approximately 4.5 metres (*Parkinson and Spiker*, 1996; *GPS performance document*, 2001).

2.2.2 The Control Segment

The control segment is responsible for maintaining the operation of the GPS. It consists of a network of monitoring stations around the world and one master station based at Colorado, United States. The monitoring stations record the raw data signal received from the satellites and send them to the master station. The master station analyses the received data to determine the satellite positions, timing errors, ionospheric delay and other corrections. These are then used to update the navigation message. The new navigation messages are transmitted to the corresponding satellites which update their broadcast signals (*GPS signal specification*, 1995).

2.2.3 The User Segment

The user segment is based around the equipment used to receive the satellite signals commonly referred to as a GPS receiver. The receivers use the satellite signals to provide the user with position information. The principle behind the receiver is to measure the delay from when a signal is transmitted from the satellite to when it reaches the receiver. This delay is proportional to the range if the velocity of the signal is assumed constant. In order to obtain the timing information, the receiver must down convert

the RF carrier wave of the signal to baseband in order to obtain the code and data. This is done by mixing the signal with a sinusoid of equal frequency to the carrier wave. However, as the distance between the receiver and satellite is constantly changing, there is a constantly changing Doppler frequency offset applied to the received signal. The receiver must track the Doppler frequency offset to mix the signal to baseband. At the same time as the carrier mixing the receiver must align a generated replica code with the code in the received signal. When the codes align they cancel leaving only the data. The amount that the replica code is shifted to achieve alignment with the signal code is used along with the pre-amble (predefined bit pattern) in the navigation data to calculate the pseudorange values (*Blewitt*, 1997).

Ideally, the range from three satellites and the satellite positions is enough to calculate the receiver position. However, in a real receiver the clock is not synchronised with the satellite clock. Satellites use extremely accurate atomic clocks which tend to be large, heavy and expensive, none of which are convenient in a receiver. Therefore a less accurate clock is used in the receiver which introduces a time error into the range measurements, known as pseudorange measurements. This timing error can be removed with an extra range measurement. Therefore four pseudorange values, combined with the precise positions of the satellites can produce a position fix. The pseudorange equations can be solved using the least-squares numerical method. A diagram of the components of a GPS receiver is shown in Figure 2.3. A detailed description of GPS can be found in *Parkinson and Spiker* (1996); *Kaplan* (1996) and the *GPS signal specification* (1995).

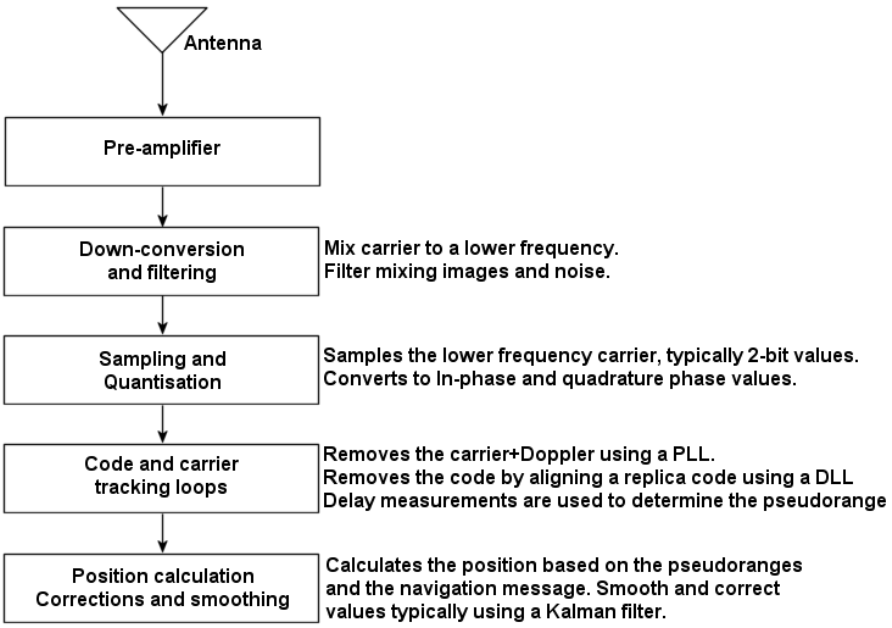


Figure 2.3: Diagram of the main components of a GPS receiver.

2.3 Signal Structure

2.3.1 Direct Sequence CDMA

Each satellite transmits a signal modulated using a form of spread-spectrum known as Direct sequence- CDMA towards the earth. By using this type of modulation all the satellites can transmit in the same frequency band. Direct sequence- CDMA works by assigning each transmitter a binary code. These codes are chosen for their low cross-correlation properties and are modulated onto a carrier using BPSK. The data is added to the signal by modulo-2-addition with the code, where the data bit period is equal to a multiple of the code period. When the signals reach the GPS receiver they have a power comparable to that of background noise. The receiver matches a replica code in phase with the received code and integrates over the code period increasing the Signal-to-Noise Ratio (SNR), known as ”de-spreading” the signal. This method was chosen so that the GPS signals do not interfere with any other ground based signals and to improve tolerance to jamming. Figure 2.4 shows a simplified example of CDMA, where the first graph has the sum of two codes representing the satellite signals. The second graph represents the replica signal generated by the receiver. The final graph shows the correlation of the two signals where a clear peak can be seen at a phase difference of zero samples.

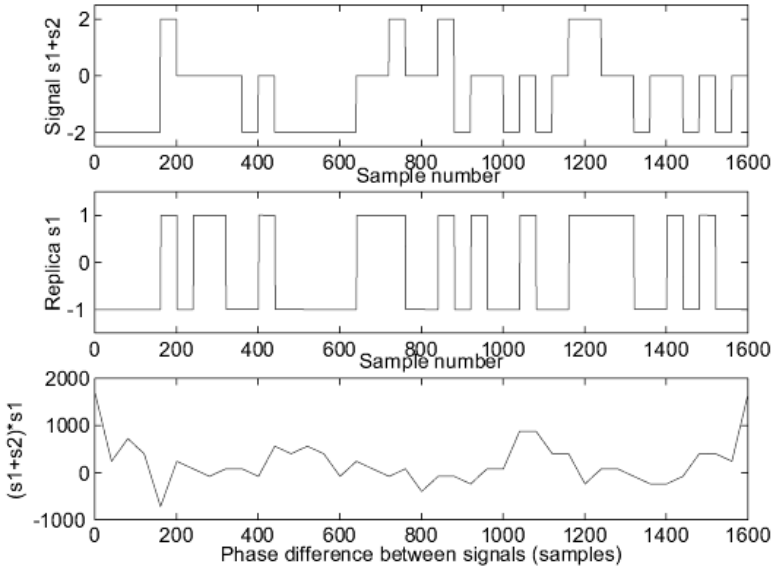


Figure 2.4: Diagram showing the principle of CDMA

2.3.2 Precise and Standard Service

The satellites broadcast on two frequency bands centred at L1 (1575.4 MHz) and L2 (1227.6 MHz). The L1 frequency consists of a code called the Coarse Acquisition (C/A) code for the SPS and a code called the Precise (P) code for the PPS. The L2 frequency is only modulated by the P code for the PPS. The C/A code is modulated in-phase (between 0 and π radian phase shifts) at a chip rate of 1.023 MHz . The P code is modulated in quadrature with the C/A code (between $\frac{\pi}{2}$ and $-\frac{\pi}{2}$ radian phase shifts) at a chip rate of 10.23 MHz . There exists the possibility to encrypt the P code, known as the anti-spoofing mode of operation, to prevent unauthorised use or jamming of the PPS. This is done by the modulo-2 summing of the P code with a lower clocked W code producing the encrypted Y code. Anti-spoofing has been enabled since the early stages of the GPS, limiting civilians to only the SPS. Another method used to degrade the SPS is known as selective availability and works by jittering the satellite's local oscillator, causing errors in the timing relating to a positional error of 100 m . The error can be removed using receivers that apply the same jitter to their clock or by differential positioning (described later). This mode was enabled for many years operation but was turned off on 2^{nd} May 2000, enabling civilian users the full accuracy of the SPS. The navigation data is added (modulo-2) to the code, where one data bit period is equal to 20 code periods (50bps). The data stream contains information on timing, corrections and satellite constellation. The structure of the navigation data can be seen in Figure 2.5. The C/A and P codes are generated by a using a shift register with feedback taps

Table 2.1: Properties of C/A and P codes

	C/A code	P code
Chip rate (MHz)	1.023	10.23
Code length (bits)	1023	$2x10^{14}$
Availability	L1	L1,L2

in order to generate a reproducible pseudo-random code. An example of the circuit used to create the C/A code is shown in Figure 2.6.

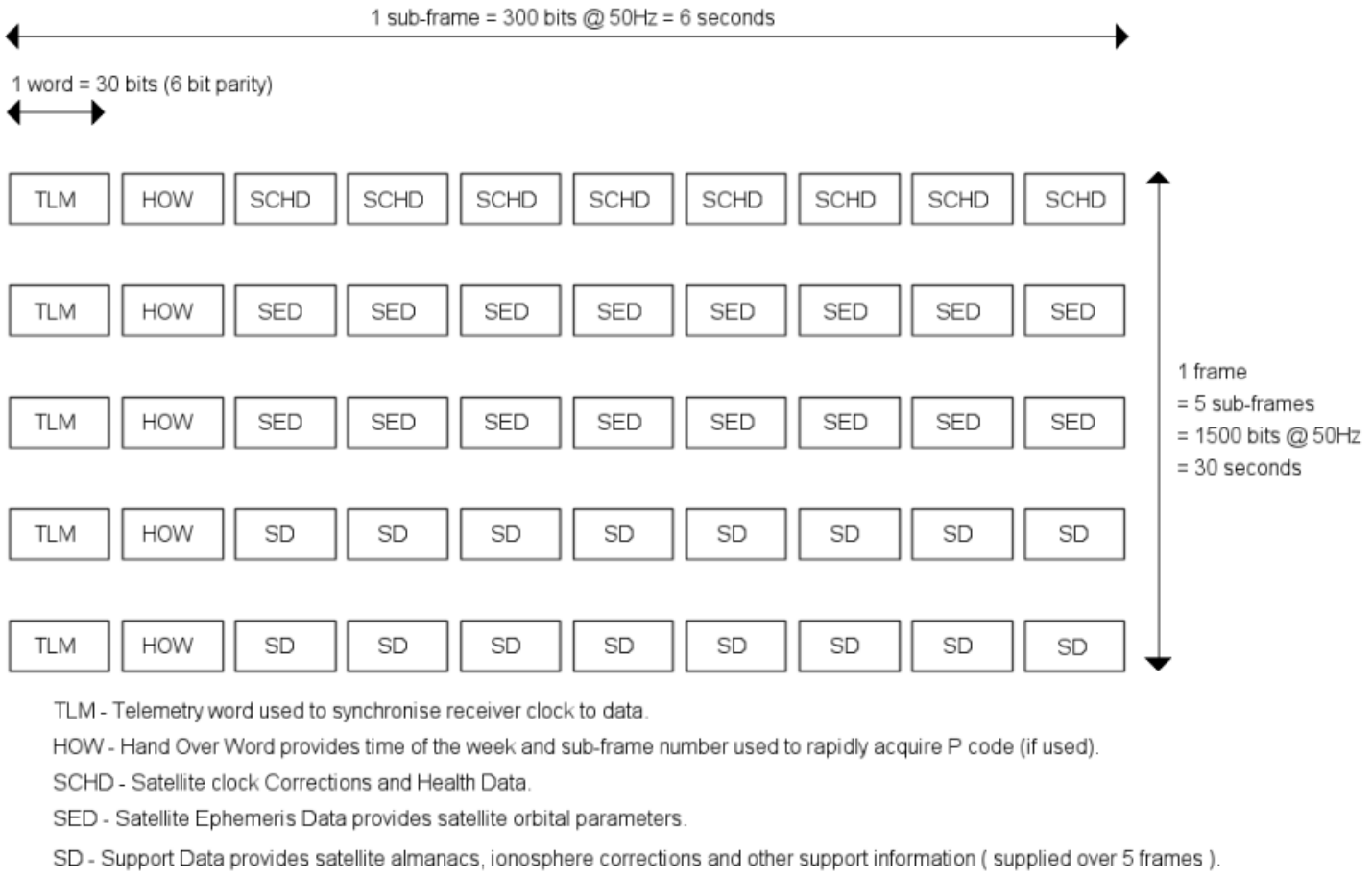


Figure 2.5: Diagram of the structure of the navigation message sent by the GPS satellites

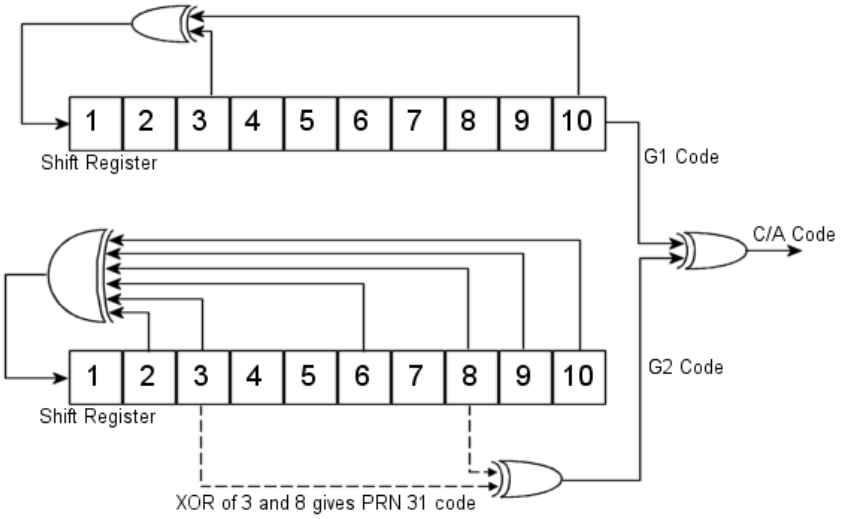


Figure 2.6: Diagram of the circuit used to generate the C/A code

There are a number of advantages in using the PPS over the SPS. The length of the P/Y code has a very low cross-correlation between signals giving the receiver more accurate error information for tracking. However, the length of the code makes it difficult to initially acquire, typically the C/A code is acquired first and the data obtained from it is used to acquire the P/Y code. The P/Y code is clocked at a faster rate than the C/A code which improves the code-tracking resolution resulting in more accurate pseudorange estimation. The P/Y code is available on both frequencies enabling a PPS receiver to calculate any delays in signal propagation due to the ionosphere. This is done by measuring the difference in propagation delay between the two frequencies, as delays due to the ionosphere are frequency dependent. Further information of the GPS signal structure can be found in, *Parkinson and Spiker* (1996); *Kaplan* (1996); *GPS signal specification* (1995).

2.4 The GPS Receiver

The GPS receiver is based around a carrier and code-tracking loop. Together they remove the C/A code and mix the carrier signal to base band, leaving only the navigation data. The delay locked loop also outputs the delay offsets required to remove the code which is used to determine the pseudorange estimate. A good overview of the processing performed in a typical receiver can be found in *Braasch and Dierendonck* (1999). A typical GPS receiver consists of an RF to IF chain, a sampling and quantisation stage, an acquisition stage, a carrier tracking stage and a code tracking stage. The following describes these stages:

2.4.1 RF to IF Chain

The GPS receiver begins with the antenna which can be based on various types. Patch antennas are the most common as they can be designed small in size, whereas helical antennas usually have better gain and LHCP rejection. This is followed by a low-noise amplifier and filters. The signal is then mixed down to an IF or close to base band frequency. It cannot be mixed directly to base band as the carrier frequency drifts due to the Doppler effect from the satellite moving in orbit. In most modern receivers the signal is digitised at this point; however some older receivers continue using the analogue signal. The principle is the same either way but any significant differences are highlighted.

2.4.2 Sampling and Quantisation

The sampling and quantisation usually occurs after the signal is down converted to a frequency close to base band. Earlier receivers tracked the carrier in the analogue domain then digitised at baseband as an Analogue to Digital Converter (ADC) of sufficient speed was not available. The sample rate and quantisation type vary from receiver to receiver and depend mainly on cost and precision required. The sample rate of the ADC must be at least double the bandwidth of the signal, however there are signal processing advantages to sampling at a higher rate. The number of levels to encode to is typically low as no amplitude information is required, however there can also be signal processing advantages in using higher levels but the cost is usually prohibitive. Currently, a typical standard civilian receiver uses 1.5bit quantisation at 5 MSps . Further information can be found in *Chang* (1977) and *Turin* (1976).

2.4.3 Acquisition

On initial start up the receiver must first acquire the satellite signals before it can track them as the signals vary in carrier and code frequency (Doppler). If the receiver has reasonably up-to-date ephemeris data it can calculate the approximate positions of the satellites and determine the likely values of the carrier frequency and code phase. The receiver can then search near these values without the need for a full acquisition. However, if there is a long period between when the receiver is turned off and back on, the ephemeris data becomes out of date and the receiver must instigate a full acquisition. The full acquisition consists of a two-dimensional search in carrier frequency and code phase. Typical receivers search +/-5 kHz from the non-Doppler carrier frequency in 500 Hz steps and through the whole code length (1023 bits, 1 ms) in 0.5 bit periods. This gives a grid of 21 carrier frequencies by 2046 code offsets, which is 42966 points to check. It is for this reason that a receiver may take some time to obtain a position

lock when first switched on. There are a number of accelerated methods for acquisition described in the next chapter.

2.4.4 Carrier-tracking

There are two options to track the carrier frequency, both are based on a Costas type of locked loop, one is the more typical Phase Locked Loop (PLL): the other is the Frequency Locked Loop (FLL) but only differs in the type of discriminator used. The FLL is the most stable as frequency changes more slowly than phase. The PLL is the most precise but is susceptible to phase noise and often cannot re-establish a lock once broken. Some receivers use a combination of the two where the FLL is used for coarse tracking and the PLL maintains a precise lock.

The Standard Phase Locked Loop

The tracking loop used in a GPS receiver is called a Costas loop, but initially a standard PLL is described. The PLL consist of a phase discriminator, a loop filter and a controllable oscillator. The PLL is described in *Kroupa* (2003); *Parkinson and Spiker* (1996).

Phase Discriminator
The purpose of the phase discriminator is to output a signal that is proportional to the phase difference between the input signals, where one input is the received signal (Equation 2.4.1) and the other is the feedback from the controllable oscillator (Equation 2.4.2) of the PLL. All analogue implementations of phase locked loops use a mixer for this purpose, however there are other types available in the digital domain. If two signals are multiplied, a term in the resulting equation is a function of the phase difference, see Equation 2.4.3.

$$S_1 = A\sin(\omega t + \phi_1) \tag{2.4.1}$$

$$S_2 = A\cos(\omega t + \phi_2) \tag{2.4.2}$$

$$S_o = A^2\sin(\phi_1 - \phi_2) + A^2\sin(2\omega t + \phi_1 + \phi_2) \tag{2.4.3}$$

Typical digital phase discriminators are usually based on similar functions to their analogue equivalents when dealing with sampled waveforms, however there are a number of functions which can track digital clocks based on exclusive-or flip-flop logic devices. These devices output a pulse train which, when averaged, is proportional to the phase difference.

Loop Filter
The loop filter removes the high frequency term from Equation 2.4.3. This leaves the sine of the phase difference which is approximately proportional to the phase differ-

ence for values between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$ radians. The transfer function of the loop filter also determines the response of the PLL by limiting its bandwidth and gain. A simple first order integrator is acceptable in many cases, however if a specific response is required higher order filters can be designed. In a digital implementation a Finite Impulse Response (FIR) or Infinite Impulse Response (IIR) digital filter can be used. The PLL is locked when the generated sinusoid is $\frac{\pi}{2}$ radians in advance of the input sinusoid and the output of the loop filter is equal to zero, see Figure 2.7. If the received sinusoid increases in frequency then the discriminator and loop filter combination produce a positive correction. This is applied to the frequency of the generated sinusoid which continually matches the received sinusoid.

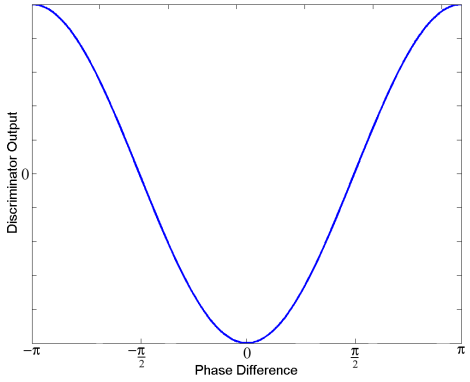


Figure 2.7: Graph showing the error correction slope of the PLL

Controllable Oscillator

In an analogue implementation this is usually a VCO which produces a sinusoid output of a frequency determined by the input voltage. The digital implementation uses a numerically controlled oscillator which is similar to the VCO but uses a discrete number of frequencies. The output from the oscillator is then fed back to the phase detector.

The Costas Phase Locked Loop and Frequency Locked Loop

The Costas PLL works by the same principle as the standard PLL and only differs in that it does not distinguish between phase differences that are π radians apart. This has a useful advantage when trying to track a binary phase modulated carrier because each data bit is modulated by π radian phase shift. A Costas loop can recover a carrier from a BPSK signal which can be used to mix the modulated signal to base band even if the carrier drifts in frequency. Figures 2.8(i) and 2.8(ii) show the difference between a standard PLL and a Costas PLL for a BPSK modulated sinusoid. At the outputs of the low pass filters (Figure 2.9, point 1) the signals are similar to the standard phase lock loop, shown in Figure 2.10 (Standard PLL). These are then fed into a discriminator

to produce a different error slope with stable points at $-\frac{\pi}{2}$ and $\frac{\pi}{2}$ radians. There are many different types of discriminator which produce different error slopes, shown in Figure 2.10 (Costas PLL). The Costas PLL discriminators are typically based on one of Equations 2.4.4 - 2.4.6.

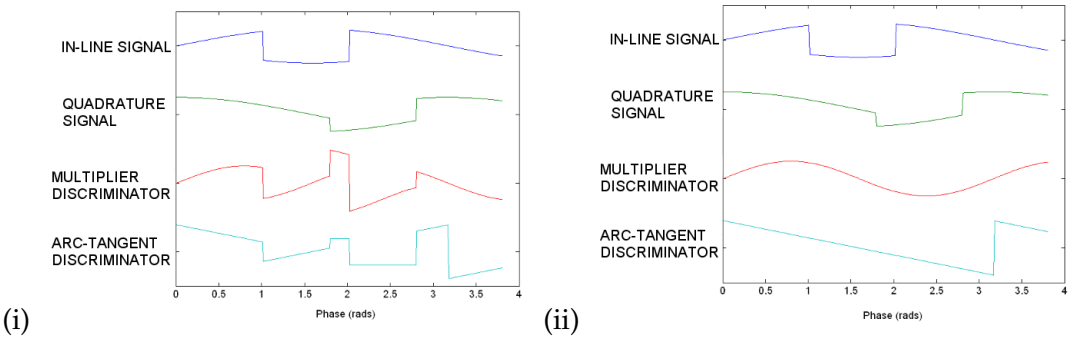


Figure 2.8: Effect of BPSK carrier on the error slope of a standard PLL (left) and a Costas PLL (right)

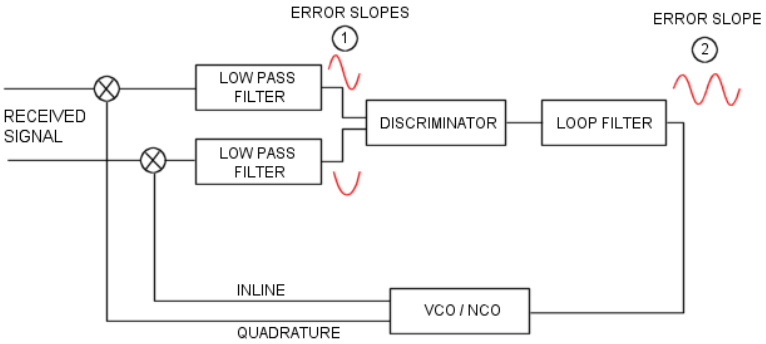


Figure 2.9: Diagram of typical Costas phase lock loop

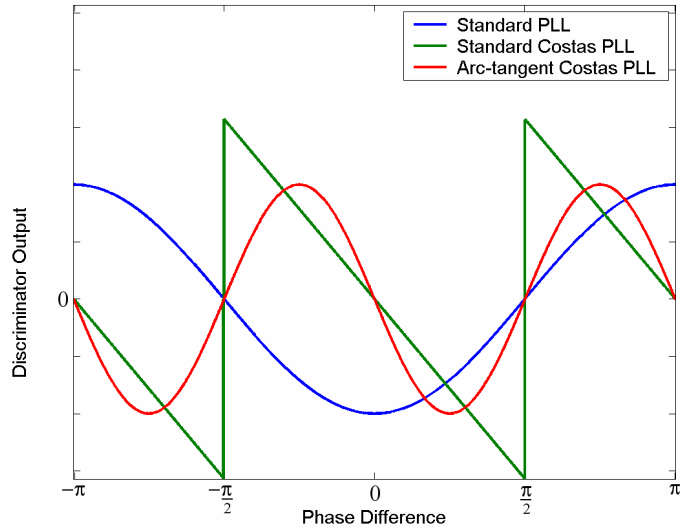


Figure 2.10: Phase difference between signals against discriminator output

Standard Costas PLL discriminator

$$\delta\phi[t] = S_Q[t]S_I[t] \tag{2.4.4}$$

Decision Directed Costas PLL discriminator

$$\delta\phi[t] = S_Q[t]sign(S_I[t]) \tag{2.4.5}$$

Arc tan Costas PLL discriminator

$$\delta\phi[t] = \tan^{-1}\left(\frac{S_Q[t]}{S_I[t]}\right) \tag{2.4.6}$$

Where,

- $sign$ is the decimal sign of the value and when tracking is equal to the navigation bit.
- $\delta\phi$ is the phase difference determined by the discriminator.
- S_Q and S_I are the integrated in-phase and quadrature signals.
- t is the sample index.

The error slopes in Figure 2.10 are output from the loop filter in the Costas loop shown in Figure 2.9 (point 2) for the different types of discriminators. Typically, the arc-tangent discriminator gives the best performance as its output is proportional to the phase difference and it is unaffected by changes in amplitude, however it is a more complex

operation than either of the others. The decision-directed discriminator tends to be less stable but provides a less noisy output when locked as the noise of the 'in-phase' channel is removed. The FLL discriminators have similar methods to the PLL discriminators but use the phase difference rather than the phase.

Standard Costas FLL discriminator

$$\delta f[t] = S_I[t-1].S_Q[t] - S_I[t].S_Q[t-1] \tag{2.4.7}$$

Decision Directed FLL discriminator

$$\delta f[t] = (S_I[t-1].S_Q[t] - S_I[t].S_Q[t-1]).sign(S_I[t-1].S_I[t] + S_Q[t].S_Q[t-1]) \tag{2.4.8}$$

Arc-tangent Costas FLL discriminator

$$\delta f[t] = \tan^{-1}\left(\frac{S_Q[t]}{S_I[t]}\right) - \tan^{-1}\left(\frac{S_Q[t-1]}{S_I[t-1]}\right) \tag{2.4.9}$$

Where,

- δf is the frequency difference determined by the discriminator.
- S_Q and S_I are the input signals to the discriminator.
- t is the sample index.

2.4.5 Code-tracking

Code-tracking is accomplished through a Delay Locked Loop (DLL) which has the purpose of aligning a received C/A or P code with an identical code generated in the loop (see Figure 2.4), thus removing the code and obtaining an offset delay for the pseudo-range calculation value. The DLL is very similar in principle to the PLL only it tracks waveforms rather than oscillations. Figure 2.11 shows the components of a typical DLL, it assumes that the carrier removal has been completed. In a real system the carrier and code removal are done in parallel but they are separated here for clarity.

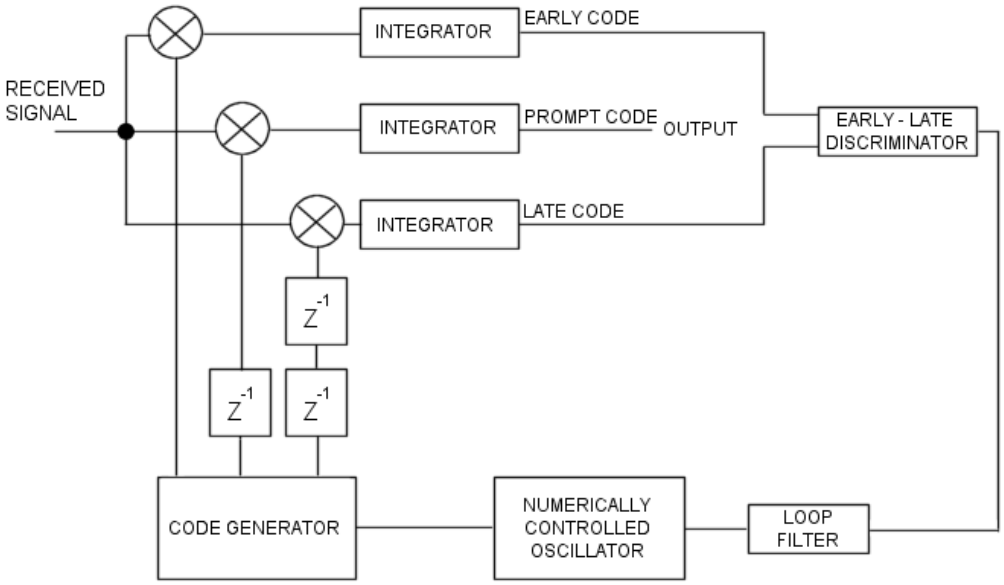


Figure 2.11: Diagram of a simplified DLL circuit

Code Generator

The code generator creates a replica code of the satellite signal. The C/A and P codes for each satellite are freely available, however access to the W codes which encrypt the P code are strictly controlled. In a typical DLL three versions of the same code are generated, each separated by half a bit period in phase. The replica codes are known as 'early', 'prompt' and 'late' codes. The DLL aims to align the 'prompt' code with the received signal where the 'early' and 'late' codes are used to form the correction signal. Many receivers have more than two sets of codes to decrease time to lock and to increase tracking robustness, (*Frank and Yakos*, 2000).

Correlator

The correlators usually consist of a mixer or an exclusive 'or' gate and an integrator. The correlator determines the correlation value between two inputs, the received signal and the generated replica code. When the received and the generated code are equal the output is one; when they are different the output is minus one. The integrator then sums these values to obtain a correlation value, which is at a peak when the codes are aligned. The integrator usually sums the values for one code period but it can be set to longer periods when tracking weak signals, providing that integration over a navigation data chip change does not occur.

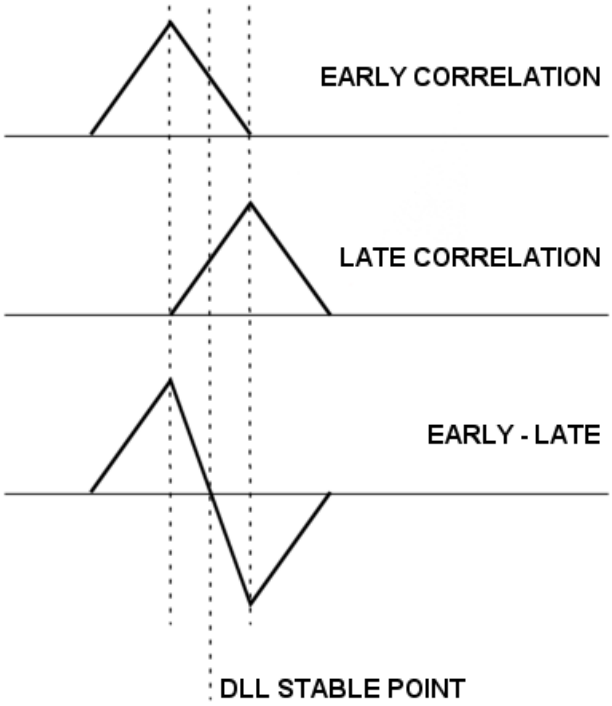


Figure 2.12: The correlation outputs and the Early-minus-late discriminator

Discriminator

There are a number of types of discriminators that can be used in a DLL but there are two main categories, coherent and non-coherent. Coherent discriminators can only track the in-phase component of the signal modulation therefore it is dependent on the PLL or FLL tracking the carrier. The non-coherent DLL can track the code even if the phase of the signal is not tracked as it combines the in-phase and quadrature components. The typical DLL discriminator is based on differencing the correlation output of two offset codes, called early and late which create an error slope, see Figure 2.12. This error slope is equal to zero when the received signal is half way between the two codes and this is maintained when the DLL is locked.

Loop Filter

The loop filter converts the error value created by the discriminator into a frequency correction for the clock of the code generator. As in the PLL the loop filter determines the response of the DLL. Many DLL's use a first order filter and rely on carrier aiding to help with the tracking. A higher order filter is required if no carrier aiding is used. Carrier aiding is a technique used to improve the code-tracking by using the more

accurate carrier Doppler corrections from the PLL. The carrier Doppler is scaled to the code Doppler and added to the code clock input, then the DLL only needs to apply small corrections.

2.5 Causes of Error in GPS.

2.5.1 The Positions of the Satellites.

The positioning used by the GPS is based on triangulation where only range values are calculated. Therefore if the satellites are close together the position is only accurate in the direction of the satellites. If the satellites are spread out then the position is accurate in each satellite direction improving the overall accuracy, see Figure 2.13. The GPS receivers use a metric called Geometric Dilution Of Precision (GDOP) to determine what constitutes good and bad satellite positions, (*Blewitt*, 1997; *Parkinson and Spiker*, 1996).

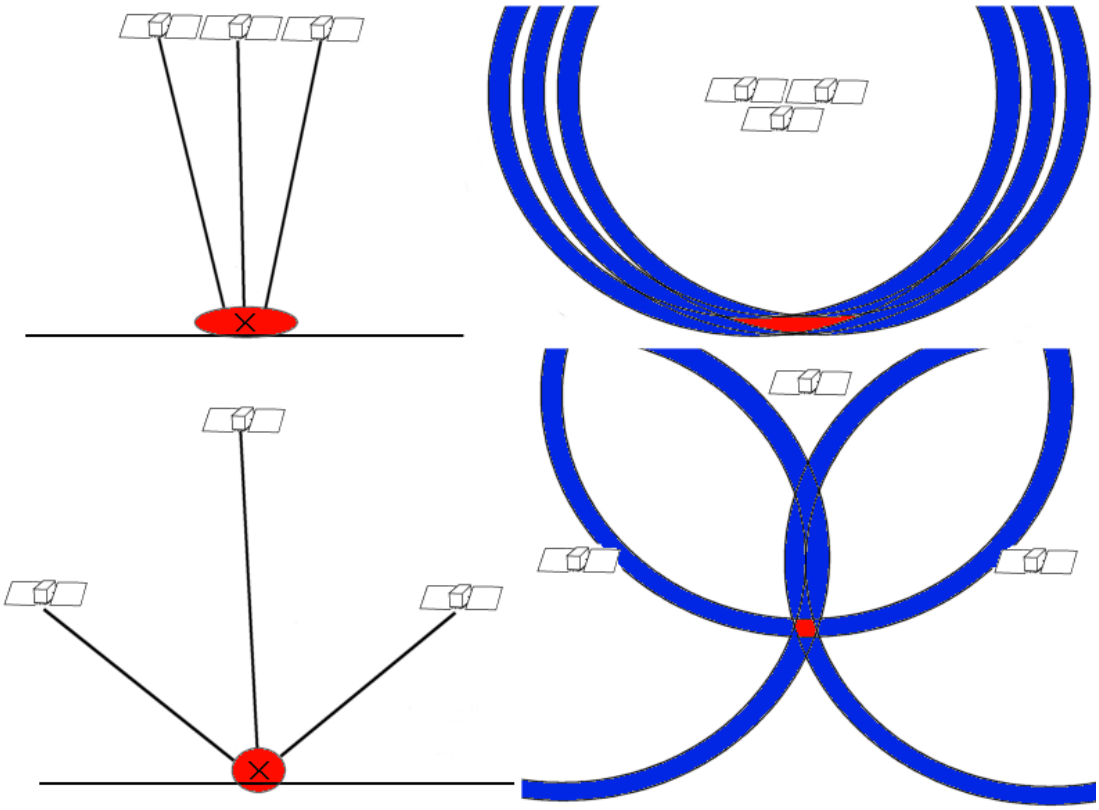


Figure 2.13: Diagram showing the effect of GDOP

2.5.2 The Atmosphere

The main regions of the atmosphere that affect the accuracy of the GPS are the ionosphere and the troposphere. The ionosphere has the strongest effect causing code group delay, carrier phase advance, scintillation and distortion of the code. The most significant errors for the GPS are the group delay and scintillation.

- Code group delay is caused by the ionosphere's refractive index which slows propagation velocity of the signal modulation. This error does not cause any problems in the receiver tracking, but does introduce an offset to the pseudorange measurements which can translate to a large positioning error.
- Carrier phase advance is also caused by the ionosphere's refractive index, but this has the opposite effect on the carrier, increasing the velocity beyond that of light. The propagation velocity of the code and carrier are of equal magnitudes from the speed of light ($V_{CARRIER} \times V_{CODE} = c^2$). The carrier travels faster than light, however it does not carry any information, no relativistic laws are broken.
- Scintillation is caused by irregularities in the ionosphere which cause diffraction and refraction of the signal. This scattering effect causes rapid changes in the amplitude and phase of the signal.
- Distortion of the code occurs because different frequencies travel at different velocities through the ionosphere, so each frequency of the modulation's spectrum is shifted in phase (*Parkinson and Spiker*, 1996). The C/A code has a relatively small bandwidth so the effects from this are negligible, however the P code has a 20 MHz bandwidth and can be affected.

An approximation of the group delay caused by the ionosphere are shown below.

$$\Delta t = \frac{40.3}{cf^2} \int N.dl \tag{2.5.1}$$

where:

- $\int N.dl$ is the known as the total electron content and is the sum of electrons in the path of the signal (el/m^2).
- f is the frequency of the carrier (Hz).
- c is the speed of light (3×10^8 ms^{-1}).
- Δt is the group delay of the signal (s).
- dl is the distance of the path through the ionosphere (m).

The properties of the ionosphere are described in Chapter 1.

At the frequencies used in satellite navigation systems (1-2 GHz) the troposphere has much lesser effect on the signals than the ionosphere. The troposphere causes a small group delay of the GPS signal which is caused by changes in the refractive index. The refractive index is dependent on the chemical composition of the troposphere and the water vapour. Delays due to the chemical composition, known as hydrostatic delays, have the strongest effect but are predictable whereas delays due to the water vapour, known as wet delays, have a smaller effect but are less predictable. An equation to estimate the zenith delay is shown below (*Bean and Thayer*, 1959), however there are many models that provide more accurate estimates summarised in *Parkinson and Spiker* (1996).

$$\Delta t = \frac{1}{c} \int N(h).dh \tag{2.5.2}$$

$$N(h) = 77.6\frac{P(h)}{T(h)} + (3.73 \times 10^{-5})\frac{e(h)}{T(h)^2} \tag{2.5.3}$$

$$(Hydrostatic) \qquad (Water\ vapour)$$

where:

- $N(h)$ is the refractivity as a function of height, which is $(n-1) \times 10^6$, where n is the refractive index
- Δt is the group delay of the signal (s).
- c is the speed of light (3×10^8 ms^{-1}).
- dh is the zenith distance (height) (m).
- $P(h)$ is the atmospheric pressure as a function of height (mb).
- $e(h)$ is the water vapour pressure as a function of height (mb).
- $T(h)$ is the absolute temperature as a function of height (K).

A weak scintillation effect is also present due to pressure irregularities and turbulence causing small-scale changes in the refractive index. The tropospheric and ionospheric errors become more pronounced at low satellite elevations as the path length increases. Most GPS receivers reject satellite signals at low elevation due to these errors.

2.5.3 The Local Environment.

The local environment describes the structures and terrain near the receiver. These can have an effect on the GPS performance due to multi-path. Multi-path is where the radio

signal finds multiple routes to the receiver via reflections, diffractions and refraction. The signal strength of the multi-path is determined by the type of interaction and the type of material. When the multiple signals reach the receiver they all have different delays due to the different lengths of each route. This causes carrier and modulation interference to the signal. The carrier interference can be constructive, where the sinusoid peaks and troughs match, or destructive, where the sinusoid peaks and troughs oppose each other. The modulation, in this case the code, is also affected but the wavelength is much longer and rather than constructive and destructive interference the code is distorted. This distortion affects the error slope of the DLL. Figure 2.14 shows an example of reflection, refraction and diffraction of a signal.

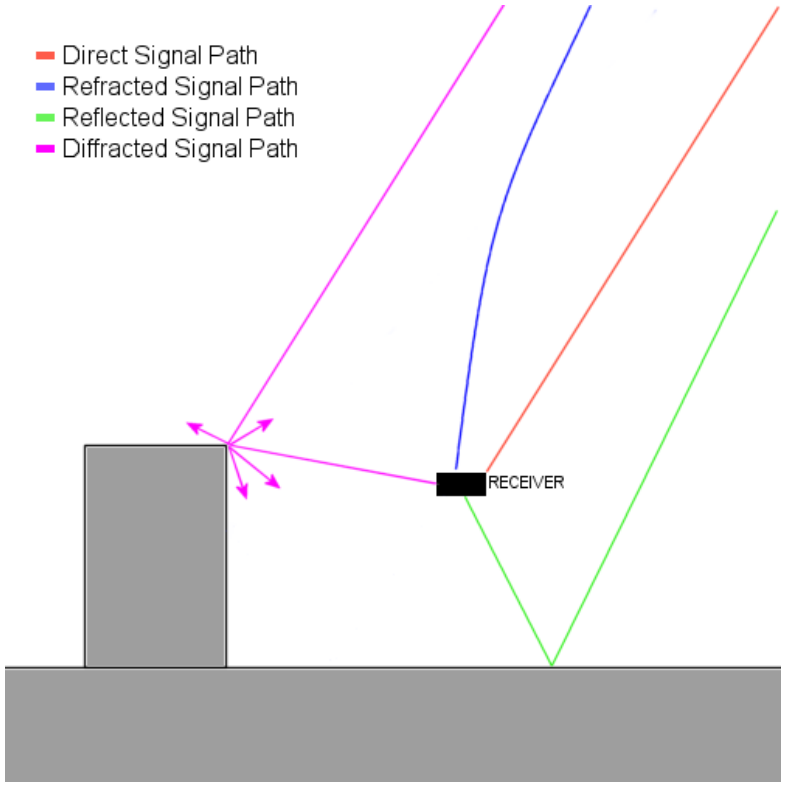


Figure 2.14: Diagram showing the possible paths a radio wave can take to a receiver

2.6 Improvements to Satellite Navigation

2.6.1 GPS Differential and WAAS

The method of differential measuring was proposed during the development of the GPS. It is based around a stationary GPS receiver of known position called the base station and a measuring receiver called a roving station. The base station receiver tracks the GPS signals similarly to the roving station except that it compares its known po-

sition with the position determined from the GPS signals. The difference between the two positions is sent to the roving receiver which is subtracted from its GPS determined position. The assumption is made that the errors due to the ionosphere are equal at the base station and roving station providing they are sufficiently close together. The factors that determine the effectiveness of the differential system are the distance between the stations, the rate of change (in distance and time) of the ionosphere and any significant multi-path effects from the surrounding terrain. The differential system is unable to offer any correction for multi-path as it varies significantly with location, this must also be considered when locating the base station as any multipath alters the corrections. The Wide Area Augmentation System (WAAS) system is a extension of the differential system where multiple corrections are broadcast to the receivers via geostationary satellites. The WAAS system covers a much larger area than the differential systems using multiple stations and interpolating algorithms to provide the corrections. The WAAS satellite covers the majority of North America. Similar systems are in development for European use (EGNOS) and Asian use (MSAS).

2.6.2 Addition of L2C and L5 GPS Signals

The L2C signal is a new GPS civilian signal to be applied to the L2 frequency band, which currently only has the restricted Y code signal. The L2C signal is a BPSK carrier modulated at the same rate as the C/A code but time multiplexed into two codes; the CM and CL codes. The CM code is 10230 chips in length and has the navigation data added to it. The CL code has no navigation data allowing variable integration times to be used in the receiver for tracking weak signals. The addition of this signal to the GPS will allow civilian users to calculate the ionospheric delays experienced by the signals in the same way as the P/Y code does for the PPS. The CM and CL codes are still clocked at the same rate as the C/A code so no significant improvement in tracking accuracy will be obtained. The first satellite with this code was launched on 26th September 2005, with three currently in orbit at time of writing.
The L5 signal is a new civilian signal designed for aeronautical use which will be broadcast on a separate frequency at 1176.45 MHz . The L5 signal is a QPSK carrier modulated with two codes, the XA and the AB, which are 10230 chips in length and clocked at 10.23 MHz . The navigation data is applied to one of the codes similarly to the L1 navigation data. With the higher code chip rate and measurements from three frequency bands (ionospheric correction) a significant improvement in tracking precision and accuracy should be seen. The first GPS satellite transmitting the L5 signal is expected to launch in 2008.

2.6.3 Development of Galileo

As well as the current satellite navigation systems, GPS and GLONASS, there is another European system in development called Galileo. Galileo is being developed as GPS and GLONASS cannot provide a guarantee of service or accuracy for the SPSs, making them unusable for safety critical systems. The system will offer a number of services 'Open', 'Commercial','Public Regulated' and 'Safety-of-Life'. The 'Open' service will provided a similar service to the current GPS SPS. The 'Commercial' service will provide access to extra ranging codes to improve performance for a fee. The 'Public Regulated' service and 'Safety-of-life' services will be restricted access services used by the police, aerospace and marine applications. These services will be particularly resistant to jamming and interference providing a guarantee of service. The fully deployed Galileo constellation will consist of 30 satellites in three orbital planes at 56 ° to the equator and at an altitude of 23000 km . The principle of the system will be much the same as GPS and GLONASS, where triangulation will be used to find the user position and a navigation message will provide the required parameters and corrections. Galileo will transmit on four frequency bands E5a (1176 MHz), E5b (1207 MHz), E6 (1279 MHz) and L1 (1575 MHz). The signal structure of Galileo has been published (*Galileo signal specification*, 2006).

- The E5a band will consist of one carrier QPSK with two binary codes, one including navigation data modulo-2 summed with the code at 25 bps , the other without data is called the pilot code. The pilot code will allow variable integration times for weak or difficult-to-track signals. The codes are 10230 chips in length and will be clocked at 10.23 MHz . This signal will provide the 'Open' service that freely available.
- The E5b band will consist of a similar signal to the E5a signal, however it will include 125 bps navigation data. This band will be used by the 'Open', 'Commercial' and 'Safety of Life' services.
- The E6 band will consist of a pilot signal and a navigation data signal, however these signals will be encrypted and only used as part of the 'Commercial' and 'Public Regulated' services. The code will be clocked at 511.5 kHz and the navigation data included at 500 bps .
- The E1 band will be shared with the current GPS L1 band and consists of a pilot and data signal. The code length is 4096 chips in length and will be clocked at 1023 kHz . The navigation message will be included at 125 bps . This band will provide signals for each of the services.

The pilot signals included in the frequency bands will use tiered codes, where the standard code is modulo-2 summed with another code (similar to the navigation data but

using set codes) allowing a receiver to use longer codes under poor signal conditions. An experimental satellite to test the signal specifications was launched in December 2005. The first set of four satellites, which will enable full testing of the system, are to be launched by the end of 2008.

2.7 Summary

This chapter has described the technology and principles of the GPS navigation system. An overview of the hardware used by the satellite and receiver has been given. The signal structure has been described including how the different services are implemented and how each satellite signal is identified. A typical receiver has also been described, which shows how the signals are used to determine the position. Some of the advancements and upcoming improvements were discussed including the European satellite navigation system which is to begin populating in 2008.

Chapter 3

GPS Receiver Technology

3.1 Introduction

Since the GPS became fully operational, work has focused on improving the accuracy and integrity of the receivers. The popularity of GPS receivers in civilian use has produced significant growth in GPS research. This chapter describes a number of the developments in GPS receiver design. The first part of the chapter describes methods that provide enhanced precision positioning using L1 and L2 carrier phase measurements. The next part of the chapter describes techniques to reduce the errors due to the ionosphere and multi-path. The final part describes a number of algorithms for rapid acquisition of the signal and an output from a typical signal acquisition is shown.

3.2 Carrier Phase Tracking

The GPS was designed to use either the C/A or P/Y codes to determine the position of the receiver. The transmission time stamp and preamble sent in the navigation message combined with the DLL delay value provides a pseudorange measurement (see Chapter 2). The problem with using the C/A code is that due to its relatively low chip rate, the correlation peak is quite wide so the DLL cannot provide as accurate a delay value as required in many circumstances. The P/Y code reduces this problem by having a chip rate that is ten times that of the C/A code, however most civilian users do not have access to the Y code. A method was called for that could take advantage of the accuracy of the carrier-tracking in the receiver, which is substantially better than the code-tracking loop. However, as their is no way to distinguish one carrier cycle from another no direct ranges can be determined. The following gives a brief overview of some of the techniques used for different applications.

3.2.1 Absolute Positioning

Many standard receivers combine the PLL and DLL outputs to improve the tracking accuracy. The aim is to obtain the robust tracking of the DLL with the accuracy of the PLL, however compromises usually have to be made in the implementation making it less than optimal. A typical method uses a Hatch filter which combines the measurements according to the following filter equation.

$$\rho(t) = \frac{(k-1)}{k}[\rho(t-1) + \delta\theta(t)] + \frac{1}{k}\tau \tag{3.2.1}$$

where:

- ρ is the corrected range measurement
- k is the filter window length
- $\delta\theta$ is the phase difference measurement from the PLL
- τ is the pseudorange measurement from the DLL

A secondary advantage of the hatch filter is its ability to reduce the effect of multi-path. As multi-path affects the pseudorange more significantly than the carrier phase, by combining them the severity is reduced. A disadvantage of this method is due to the effect of the ionosphere on the code and carrier of the signal, where the ionosphere delays the code and advances the carrier phase. Therefore under strong ionospheric conditions and depending on the window size of the filter the two separate measurements may diverge. The principle of this method has been adapted to use the weighted Hatch filter, the Kalman filter and the extended Kalman filter (*Chen et al.*, 2000), (*Dai et al.*, 2000), (*Hwang and Brown*, 1990). The Hatch filter method has also been adapted from the range domain to a position domain solution giving the advantages of stability during satellite changes (*Lee and Rizos*, 2003).

3.2.2 Relative Positioning

This type of measurement is usually only used in surveying as it can only provide a position relative to another receiver's position. The advantage of the technique is a very high accuracy, however long dwell times are often needed to achieve it. Two receivers are used, one of known position termed the static station and one to measure termed the roaming station. The technique measures the change in phase between the two receivers and a number of satellites. The carrier phase observations from a receiver can be described by:

$$\phi = \rho + c(\delta T_s - \delta T_r) + \lambda N + \epsilon \tag{3.2.2}$$

where:

- ρ is the true range
- c is the speed of light
- T_s is the timing bias error from the satellite clock
- T_r is the timing bias error from the receiver clock
- λ is the carrier wavelength
- N is the carrier cycle ambiguity
- ϵ describes the bias errors from other sources (ionospheric and tropospheric delays)

Many of the bias errors shown above can be eliminated by differencing simultaneous observations. A typical double-difference method can remove the timing bias errors leaving only the carrier cycle ambiguity (see Figure 3.1). The method differences observations between two receivers from a number of satellites at the same time. Then these differences are differenced again between twin satellites. This method also converts the ambiguity, N into an integer value.

$$\phi_{dd} = \rho_{dd} + \lambda N_{dd} + \epsilon_{dd} \tag{3.2.3}$$

The equation can be converted to a least squares form in order to solve the carrier cycle ambiguity. The solution assumes that the error bias due to the atmosphere is equal at both receivers.

$$\Gamma = \phi_{dd} - \rho_{dd} - \lambda N_{dd} \tag{3.2.4}$$

where Γ is the residual. When a number of double differenced observations are obtained the least squares equation can be solved.

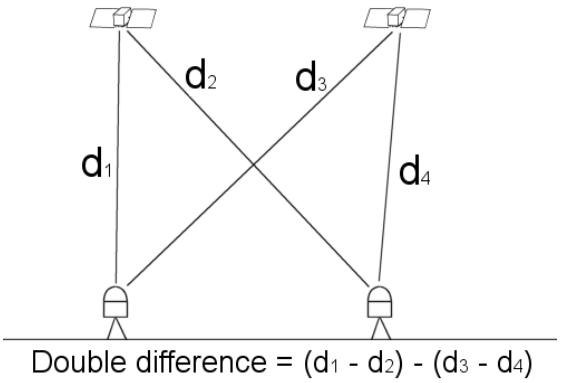


Figure 3.1: The double difference method

Further developments including the benefits of dual frequency tracking and other methods can be found in *Verhagen* (2005); *Moon and Verhagen* (2006); *Teunissen and Verhagen* (2007).

3.3 Dual frequency operation

As described in Chapter 2, the GPS satellites transmit two frequencies, L1 and L2, but the C/A code is only available on L1, meaning L2 is unavailable for the standard positioning service. The L2 code is modulated by a code known as the P code. The P code is publicly available, however the GPS operates under a mode known as anti-spoofing, where the P code is encrypted by combining it with the W code, forming the Y code. The Y code is only used by the US DoD receivers and access to the W code is strictly controlled. The reason for this anti-spoofing mode is to prevent the possibility of false GPS signals causing receivers to read incorrectly in critical situations. As well as this protection, the other advantages of using the Y code are more precise pseudorange calculation and ionospheric delay corrections. Despite no access to the W code, it can still be possible to determine the ionospheric delay from the L2 frequency using a number of techniques which vary in success. An overview of current methods and suggested new methods can be found in *Woo* (1999).

3.3.1 Squaring

This method enables the carrier tracking of the L2 frequency, but no pseudorange can be determined so it can only be used for relative positioning. The method multiplies the signal with itself, squaring the signal. This has the effect of removing all of the modulation (BPSK) and doubling the frequency of the carrier. Unfortunately, before despreading the signal is typically of lower power than the background noise, squaring increases the noise power substantially. This causes significant tracking problems for the phase lock loop in the receiver making the Doppler calculation error prone. Another effect of the squaring is a doubling of the carrier frequency which makes the carrier cycle ambiguity resolution more difficult (12 cm steps). The PLL must also resist tracking different satellite signals on L2, as there is no code to distinguish between them, any signals with similar Doppler shifts can be confused. This also makes identifying the source of the transmission difficult in some cases, which is done by comparing the Doppler values with those from the signals identified on L1. The squaring technique was used in early GPS receivers but is rarely used now due to better methods.

3.3.2 Code aided squaring

This method enables carrier phase and pseudorange measurements from L2, allowing the determination of ionospheric delays. The method exploits the similarity of the freely available P code to the encrypted Y code. As mentioned earlier in this section the P code is encrypted by multiplying with a lower clocked W code, therefore if the Y code is mixed with the P code the lower clocked W code remains. As the bandwidth of the W code (1 MHz) is much lower than the Y code(20 MHz), it is possible to remove more of the background noise by filtering. The P code is generated very similarly to the C/A code (shown in Chapter 2) at early, prompt and late offsets, which are fed into a DLL. The remaining W code is removed using squaring, resulting in a carrier signal of half wavelength but of higher signal to noise ratio than the squaring only method.

3.3.3 L1 and L2 cross correlation

This method also enables carrier phase measurements and a form of pseudorange measurements from L2. The method exploits the fact that the Y code is modulated on both L1 and L2 frequencies. The C/A code is removed from the L1 frequency leaving only the Y code on both L1 and L2 frequencies. Although the Y codes are transmitted from the satellite at the same time on both frequencies, the ionospheric delay is different at each frequency misaligning them. Therefore the codes must be realigned at the receiver by introducing a delay to one of the signals. The delay calculated can be used to determine the ionospheric delay of both frequencies using Equation 3.5.1. The L1 and L2 signals are then mixed together removing the Y code and producing a carrier at L1-L2 frequency. The lower frequency can be used for solving the carrier cycle ambiguity in relative positioning as the wavelength is 86cm (much greater than L1). To obtain the L2 carrier the L1-L2 carrier can be mixed with the L1 carrier. Unfortunately as the multiplying is done at the full bandwidth of the Y code, the background noise is severe. The signal to noise is better than the squaring method as the L1 signal is transmitted at a greater power than L2. More detailed information on this technique can be found in *MacDoran et al.* (1985).

3.4 Multi-path Mitigation

Multi-path has been a problem in the GPS since its implementation, however as the use of GPS moves into more urban environments multi-path reflections become a more severe source of error. A number of solutions to reduce the problem have been implemented. Until recently, these solutions have focused on blocking signals from certain arrival elevation angles and have relied upon antenna design.

3.4.1 Antenna Design

Many current antennas are designed to attenuate any signals that arrive below 10 degrees in elevation. They rely on the fact that multi-path usually arrives at the receiver from below that elevation and that the antenna remains approximately vertical. This is becoming a less valid assumption with receivers becoming more portable and being used in urban environments. A more recent solution is to use antenna arrays to form a beam that tracks the satellites independently (*Brown et al.*, 2001). This technology has been used extensively in radar applications and therefore many innovations have already been implemented like adaptive nulling and beam shaping. Some of these innovations could be useful in GPS applications however the significant processing requirements still make these impractical in most situations. Antenna arrays can provide very good mitigation of multi-path, but are currently impractical in most situations due to the size of the array and the amount of processing required.

3.4.2 DSP Techniques

As described in the carrier-tracking section, there are methods that use carrier smoothing to reduce multi-path effects. This is possible because the PLL tracks on a much smaller scale than the DLL. For example a 10% error in the DLL discriminator causes a 30 m range error, where as a 10% error in the PLL discriminator only causes 2 cm range error. These methods are widely used, however the improvement is modest and there are a number of drawbacks depending on the method of combining the measurements. Another method widely used is the narrow correlator method where the early and late codes used by the DLL are separated by a smaller delay, typically one fifth of a chip. This requires a higher sample rate at the ADC of the receiver IF chain, approximately ten times that of the C/A code. By narrowing the discriminators error slope, it becomes much less affected by nearby multi-path. Figure 3.2 shows the error slopes of the standard correlator DLL on the left and the narrow correlator DLL on the right with multipath (green) and without multi-path (blue). It can be seen that the error slopes of the narrow correlator with multi-path are much closer to the non multi-path cases. A performance evaluation and more details can be found in *Dierendonck et al.* (1992) and *Braasch* (2001).

The Multi-path Estimating Delay Lock Loop (MEDLL) is a much more active approach to multipath. It uses a large number of correlators to determine the amplitude, code delay and carrier phase of the direct and multi-path signals. This is done by using maximum likelihood or least squares estimation to find the set of direct and multi-path signals that reproduce the correlation pattern observed. The direct ray can be identified as it is the earliest and strongest signal received. Once the properties of the signals are established their effect can be subtracted from the measured correlation function leav-

ing an undistorted correlation function. This approach requires a significant increase in processing and hardware but has achieved much success. The description can be found in *van Nee* (1992) and implementation and evaluation information can be found in *Townsend et al.* (2000)

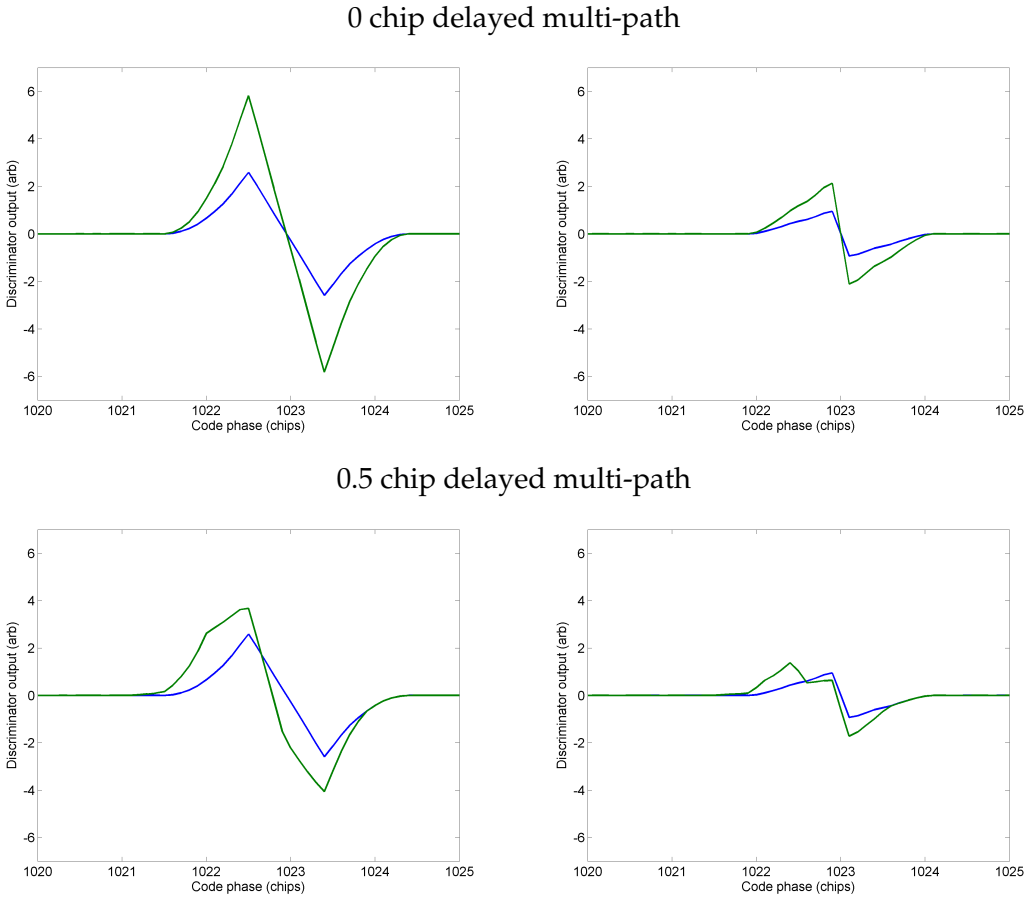


Figure 3.2: Figures of the standard (left) and narrow correlator (right) error slope with half power multi-path at half and one chip offsets

3.5 Ionosphere Corrections

3.5.1 Group Delay

The SPS provides coefficients for an ionosphere model as part of the navigation message which can reduce the bias error by 50%. This may be acceptable in some applications however, many require much greater accuracy. The PPS tracks the P/Y code on two frequencies enabling the ionospheric bias error to be determined to a 1-2 m accuracy. The ionosphere correction can be calculated using the cross-correlation or semi-codeless techniques described in the 'Dual Frequency' section earlier in this chap-

ter. The ionospheric delay is calculated from the difference in the arrival times of the L1 and L2 signals shown in Equation 3.5.1. This method is significantly less accurate than the PPS method however has much improved accuracy on the SPS ionosphere model. Other methods rely on differential techniques, where a receiver of known location transmits the position errors to receivers nearby (see Chapter 2).

$$\Delta t_1 = \frac{f_2^2}{f_1^2 - f_2^2} \delta t_{1-2} \tag{3.5.1}$$

where:

- Δt_1 is the ionospheric delay for the L1 frequency
- f_1 is the L1 frequency
- f_2 is the L2 frequency
- δt_{1-2} is the difference between the L1 and L2 arrival time

3.5.2 Scintillation

Scintillation is the rapid change in phase and amplitude of the signal, which cause the tracking loops to fluctuate and even lose lock in some circumstances. Widening the bandwidth of the tracking loops reduces the chance of a loss of lock but increases the noise in the range measurements, where as narrowing the bandwidth causes the opposite effect. If too wide a bandwidth is used the discriminator error slope degrades into noise making tracking impossible. *Humphreys et al.* (2005) developed a Kalman based tracking loop which effectively varies the bandwidth based on the Carrier to Noise density (C/N_0) of the signal enabling it to track under significant scintillation. Other methods are mainly in the position domain (rather than the range-domain) and treat scintillation the same as background noise using standard tracking filters like the Kalman filter and more recently the particle filter, (*Kao and Eller*, 1983; *Gustafsson et al.*, 2002).

3.6 Signal Acquisition Techniques

3.6.1 The Standard Acquisition

There are three main methods used to acquire the GPS signals; the standard search method, the carrier frequency based FFT method, the code delay based FFT method. The first method shown in Figure 3.3 searches through the code phases and carrier frequencies until a peak in the correlation is obtained. The search consists of multiplying

the signal by a replica code and carrier which are varied in phase and frequency respectively. The typical search space consists of 500 Hz steps, 5kHz either side of the transmitted carrier frequency and the full code range in half chip steps. This gives a search space of 42966 bins, which can take up to 5 minutes to do on a typical receiver.

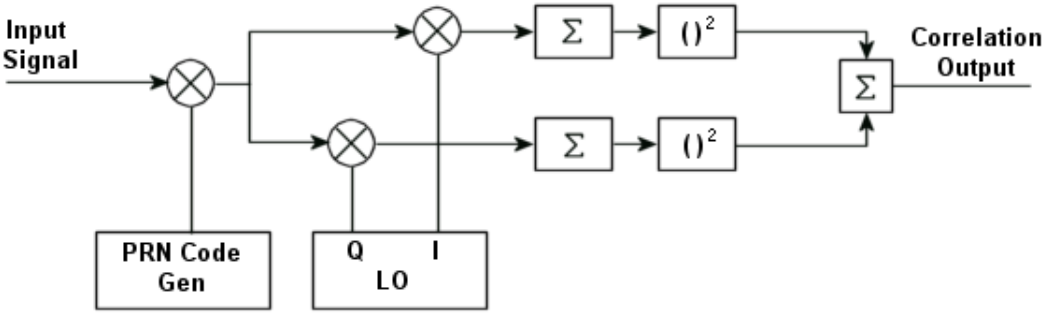


Figure 3.3: Diagram of the standard acquisition process

3.6.2 The Carrier Frequency Accelerated FFT method

The second method shown in Figure 3.4 performs a Fast Fourier Transform (FFT) for every half chip of the C/A code. As the FFT cannot be limited to a range of frequencies, a full frequency spectrum of the signal is taken at each code phase. Despite this problem, the method is moderately better than the full search method providing the receiver hardware is capable of performing the function.

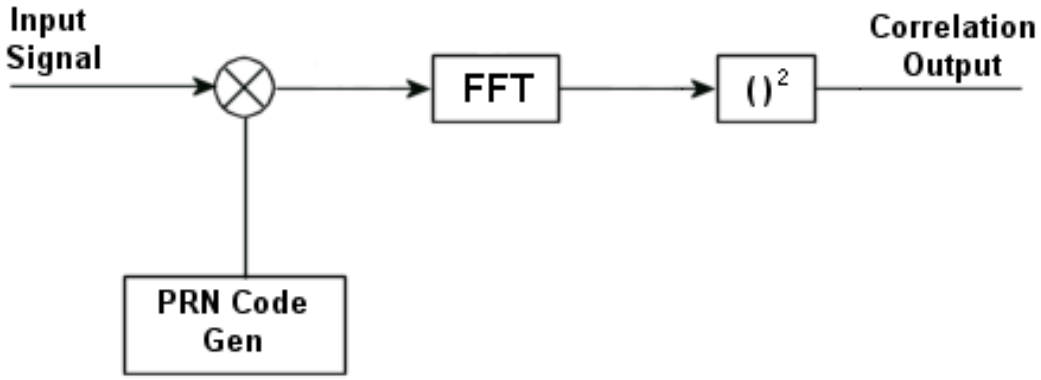


Figure 3.4: Diagram of the frequency accelerated process

3.6.3 The Code Accelerated FFT method

The third method shown in Figure 3.5 searches the carrier frequency in the same way as the standard method. However, the code is reversed and convoluted with the mixed signal. This has the effect of correlating the replica code with the signal but using the efficiency of the FFT. This method reduces the code search time, which is the longest stage of the acquisition so even with two FFTs and an inverse FFT, it is still faster than the other methods.

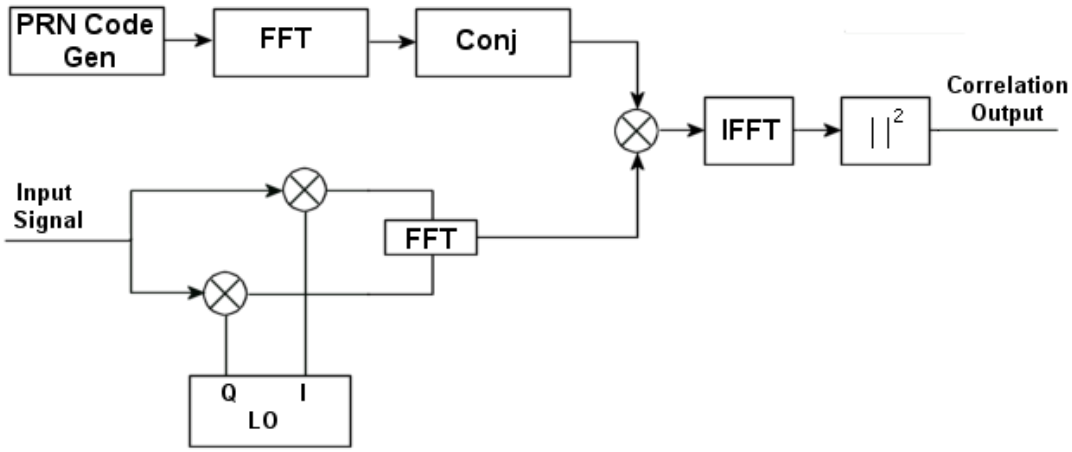


Figure 3.5: Diagram of the code accelerated process

Ward (1996) shows a number of the search techniques used in receivers and *Djebouri et al.* (2004) shows the enhanced search algorithm using the FFT. Figure 3.6 shows the output from this type of acquisition used in the simulator.

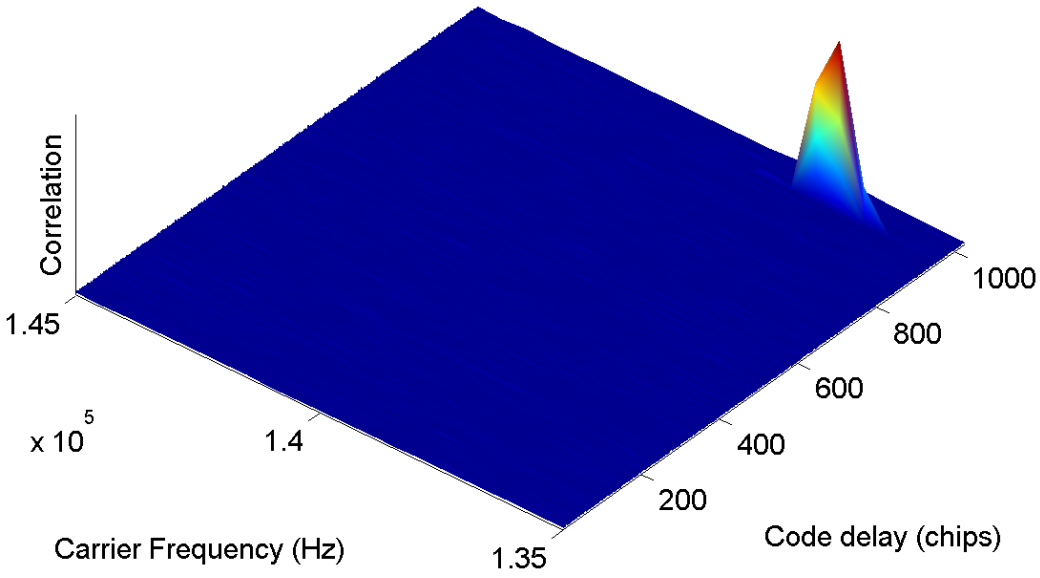


Figure 3.6: Output from the acquisition function of the simulator

3.7 Summary

This chapter has described the main areas of GPS receiver research. The start of the chapter described carrier tracking methods for L1 and L2 frequencies, which can be used for relative positioning methods (surveying) or to improve the precision of the pseudorange measurements using smoothing algorithms. A summary of the methods to correct the errors due to multi-path and the ionosphere was provided. The chapter concluded with descriptions of some improved signal acquisition methods and an output from a typical acquisition is shown.

Chapter 4

Simulation of GPS

4.1 Introduction

This chapter describes the GPS simulator that has been developed for this project, from the reason why certain models and methods were used to the assumptions and approximations made. Weighing the accuracy of the simulation against the computational complexity and memory available was a major part of this project. The chapter begins with descriptions of general radio propagation models followed by a discussion and a selection of a model. This is followed by a summary of the ionosphere delay and scintillation models. The final section describes how the simulation works and how the models were integrated.

4.2 General Radio Propagation Model

Fundamental to the simulation of the GPS is how to model the actual propagation of the signal. This initial model is then adapted to include the ionospheric and multipath errors in the GPS. This section introduces a number of the propagation models available and then discuss the most suitable for the simulation.

4.2.1 The Parabolic Wave Equation Model

The Parabolic Wave Equation model (PWE) is derived from the Helmholtz wave equation and is a forward scatter narrow angle approximation. A general case for the PWE is described below. More detailed descriptions of various implementations of the PWE approximation can be found in *Levy* (2000). An implementation of the PWE model for GPS can be found in *Hannah* (2001).

The scalar wave equation below describes propagation of a wave, ϕ, through a medium of constant or slow changing refractive index, n, in a 2-D environment (x,z). It assumes that the wave is of the form $e^{-j\omega t}$,where ω is the angular frequency and t is the time and it has a wave number, k.

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial z^2} + k^2 n^2 \phi = 0 \tag{4.2.1}$$

If the assumption is made that the wave is planner and travels in the x-axis direction.

However, to simplify the equation the assumption is made that the signal varies slowly in the x direction. The wave can then be described by:

$$u(x, z) = \phi(x, z) \exp^{-jkx} \tag{4.2.2}$$

Substituting u into the scalar equation forms the following equation.

$$\frac{\partial^2 u}{\partial x^2} + 2jk\frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial z^2} + k^2 u(n^2 - 1) = 0 \tag{4.2.3}$$

As the signal varies slowly with increasing x the term $\frac{\partial^2 u}{\partial x^2}$ can be assumed negligible.

The remaining equation can be split into forward and backward scattering terms for the wave.

$$\left\{\frac{\partial u}{\partial x} + jku(1 + Q)\right\}\left\{\frac{\partial u}{\partial x} + jku(1 - Q)\right\} = 0 \tag{4.2.4}$$

where,

$$Q = \sqrt{\frac{1}{k^2}\frac{\partial^2}{\partial z^2} + n^2} \tag{4.2.5}$$

The angle from the x-axis at which the model errors become large is determined by the accuracy of the Q approximation. Q contains the second order differential of the wave with respect to the z-axis which determines the propagation of the wave in the z-axis direction. The standard PWE approximates the square-root in equation 4.2.5 as a first order Taylor expansion, which limits its propagation angle to less than approximately 15 $^\circ$ from the x-axis direction.

The parabolic equations can be solved in a number of ways, however the two most common are the Fourier split-step technique and the implicit finite-difference technique. The Fourier split-step technique provides the most computationally efficient method of solution, however the implicit finite-difference technique can accept more complex boundary conditions. Descriptions of this techniques are shown in *Kuttler* (1999); *Brent and Ormsby* (1995); *Hyaric* (2001). The PWE provides a computationally efficient solution for a wave method.

4.2.2 Finite Difference-Time Domain

The FD-TD model is simple in essence, but applying it to detailed three dimensional environments is both complex and computationally intensive. It approximates Maxwell's curl equations using difference operators instead of differentials. The model divides the volume of interest into equal sized cubic segments, where each cube must be small in comparison to the wavelength of the electromagnetic field. It then solves the equations for each cube based on the values of its neighbours. Implementations of FD-TD models can be found in *Lee* (2003) and *Buchanan* (1996).

The Maxwell curl equations for an isotropic medium,

$$\nabla \times E = -\frac{\delta B}{\delta t} \tag{4.2.6}$$

$$\nabla \times B = \mu_0 J + \mu_0 \epsilon_0 \frac{\delta E}{\delta t} \tag{4.2.7}$$

where,

- B is the magnetic field strength (T).
- E is the electric field strength (Vm^{-1}).
- J is the current density (Am^{-2}).
- t is the time (s).
- μ_0 is the magnetic permeability of the medium (NA^2).
- ϵ_0 is the electric permittivity of the medium (Fm^{-1}).

The vector equations can be split into the scalar components and applied to the environment.

$$\frac{\partial B_x}{\partial t} = \frac{1}{\mu}\left(\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y}\right) \tag{4.2.8}$$

$$\frac{\partial B_y}{\partial t} = \frac{1}{\mu}\left(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z}\right) \tag{4.2.9}$$

$$\frac{\partial B_z}{\partial t} = \frac{1}{\mu}\left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x}\right) \tag{4.2.10}$$

$$\frac{\partial E_x}{\partial t} = \frac{1}{\varepsilon}\left(\frac{\partial H_z}{\partial y} - \frac{\partial B_y}{\partial z} - \sigma E_x\right) \tag{4.2.11}$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\varepsilon}\left(\frac{\partial H_x}{\partial z} - \frac{\partial B_z}{\partial x} - \sigma E_y\right) \tag{4.2.12}$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon}\left(\frac{\partial H_y}{\partial x} - \frac{\partial B_x}{\partial y} - \sigma E_z\right) \tag{4.2.13}$$

The 'Yee' FDTD method (*Yee*, 1966) is described below. The calculations are applied to the environment in the form of cubes shown in Figure 4.1.

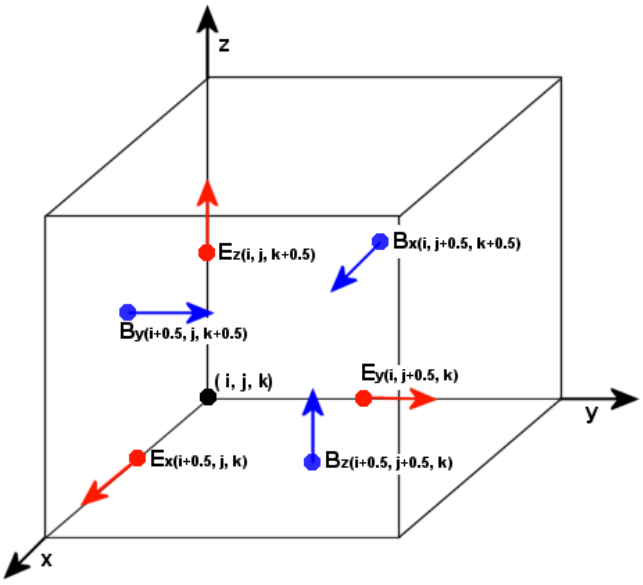


Figure 4.1: Diagram the solution grid for Yee's FDTD method

The first order differential terms are approximated as,

$$\frac{\partial B_x}{\partial t} \approx \frac{B_x(t+\Delta t) - B_x(t)}{\Delta t} \tag{4.2.14}$$

$$\frac{\partial B_x}{\partial y} \approx \frac{B_{x+\frac{\Delta x}{2}}(t) - B_{x-\frac{\Delta x}{2}}(t)}{\Delta y} \tag{4.2.15}$$

where x and y can be substituted for any axis direction.

These difference terms can be substituted into the scalar equations to form an algorithm in the discrete domain. For the x-axis components:

$$B_{x^{(a,b+0.5,c+0.5,n+0.5)}} - B_{x^{(a,b+0.5,c+0.5,n-0.5)}} = \frac{E_{z^{(a,b+1,c+0.5,n)}} - E_{z^{(a,b,c+0.5,n)}}}{\Delta y} - \frac{E_{y^{(a,b+0.5,c+1,n)}} - E_{y^{(a,b+0.5,c,n)}}}{\Delta z} \tag{4.2.16}$$

$$E_{x^{(a+0.5,b,c,n+1)}} - E_{x^{(a+0.5,b,c,n)}} = \frac{B_{y^{(a+0.5,b,c+0.5,n+0.5)}} - B_{y^{(a+0.5,b,c-0.5,n+0.5)}}}{\Delta z} - \frac{B_{z^{(a+0.5,b+0.5,c,n+0.5)}} - B_{z^{(a+0.5,b-0.5,c,n+0.5)}}}{\Delta y} \tag{4.2.17}$$

where,

- a is the current position on the x-axis
- b is the current position on the y-axis
- c is the current position on the z-axis
- n is the current sample in time

It can be seen that each field value is calculated by the value of the neighbouring alternate field values.

4.2.3 Ray-shooting Model

The ray-shooting model is based on the theory of geometric optics which describes the behaviour of electromagnetic waves by using a number of straight lines called rays. The model is based around a ray-shooting algorithm which calculates ray paths and their interactions with obstacles. The ray assumption can apply to any frequency electromagnetic wave, but only reflections with objects of larger sizes than the wavelength of the wave are modelled accurately. The model uses the Friis free space path loss formula to calculate the power loss from propagation.

$$P_{RX} = \frac{P_{TX} G_{TX} G_{RX} \lambda^2}{16\pi^2 d^2 L} \tag{4.2.18}$$

where

- P_{RX} is the power received.
- P_{TX} is the power transmitted.
- G_{TX} is the antenna gain of the transmitter.
- G_{RX} is the antenna gain of the receiver.
- λ is the signal wavelength.
- d is the propagation distance.
- L is the factor representing system losses.

Reflection

Power loss from reflections are determined using squared Fresnel reflection coefficients and multiplying with the propagation losses. The modified equation becomes:

$$P_{RX} = \frac{P_{TX} G_{TX} G_{RX} \lambda^2 \Gamma^2}{16\pi^2 d^2 L} \tag{4.2.19}$$

where the Fresnel reflection coefficient, Γ is determined by the following equations for a horizontally and vertically polarised signal respectively:

$$\Gamma_H = \frac{sin\theta - \sqrt{\epsilon - cos^2\theta}}{sin\theta + \sqrt{\epsilon - cos^2\theta}} \tag{4.2.20}$$

$$\Gamma_V = \frac{\epsilon sin\theta - \sqrt{\epsilon - cos^2\theta}}{\epsilon sin\theta + \sqrt{\epsilon - cos^2\theta}} \tag{4.2.21}$$

4.2.4 Physical Optics Model

The physical optics model uses ray optics to determine which parts of the surface are illuminated by the field, then it calculates the currents induced in the surface by this field. The scattered field is determined by integrating the calculated currents. The assumption is made that all surfaces are perfect conductors and are infinite planes. The following describes a simplified 2-Dimensional version of the model described fully in *Burke* (1994).

The first step of the Physical Optics (PO) model is to evaluate the surface current density at areas on the surface, as a result of the incident wave. For illuminated parts of the surface, assuming a perfect reflection,

$$J_{Surf} = 2\hat{n}\vec{H}_i$$

For shaded parts,

$$J_{Surf} = 0$$

Where $\hat{n}$ is a unit vector perpendicular to the reflective plane and $\vec{H}_i$ is the incident magnetic field. Setting the shadowed area currents to zero reduces the accuracy of the model but simplifies the solution.

The equation for the scattered field is the integral of the currents formed in the surface by the incident fields,

$$E_{Scatter} = -j\omega\mu \int_S G(\hat{r_e}, \hat{r_s}) J_{Surf}(\hat{r_s}).\delta S \tag{4.2.22}$$

Where,

- $\hat{r_s}$ is the vector to the source point
- $\hat{r_e}$ is the vector to the field evaluation point
- S is the surface area
- J_{Surf} is the current density on the surface

- $G(\hat{r_e}, \hat{r_s})$ is the dyadic Green's function
- $E_{Scatter}$ is the electric field scattered from the surface
- μ is the permeability of the surface
- ω is the angular frequency of the field

The PO model is only accurate when the dimensions of the scatterer are large in comparison to the wavelength of the incident field. This is due to the assumption made when calculating the surface current density that the reflection plane is infinitely long. This assumption also causes the reflections near plane edges and diffraction effects to be incorrectly modelled. An implementation of the model for rough surface scatter can be found in *Burklioltler* (2003); *Zalin and Sarahandi* (1999); *Burke* (1994). A modified version including diffraction modelling can be found in *Umul* (2004).

4.2.5 Conclusion

Table 4.1: Properties of radio propagation models

	Accuracy	Computational complexity	Typical Environment	Signal model
PWE	Very good (wave-based)	Moderate	2D	Carrier, no modulation
FD-TD	Excellent (wave-based)	High (especially for 3D)	2D or 3D	Carrier, no modulation
Ray-shoot Model	Moderate (ray-based)	Low	3D	Waveform possible
PO Model	Moderate (ray-based)	Moderate	2D or 3D	Waveform possible

Table 4.2.5 shows a comparison of the propagation models investigated. The most important factor for GPS propagation model is how accurately the phase of the signal is modelled with the accuracy of the intensity second. The GPS relies on very accurate timing of the signals to operate, where as most receivers quantise the signal intensity to two or less bits. The intensity model is still important as it is accumulated to form the correlation; however lower accuracy values should not cause significant errors to the simulation. From the models above most model intensity and phase information however the FDTD and the PWE are limited to only carrier wave models. The ray-based method produces accurate phase information and reasonable intensity accuracy, although it does not model the wave aspects of the signal (diffraction and refraction). The density of rays transmitted does affect the accuracy but this can be scaled easily with the speed of execution. Many of the other propagation models discussed require

substantial computational resources especially for the large environment GPS covers. This must be considered as computational power is limited and there are a number of other parts to the simulation. From the analysis, a ray-shooting Geometric Optics (GO) based model using an algorithm common to the computer graphics community was developed for use in the simulation. The algorithm used is based on the 3 dimensional intersection of a line with a plane described below.

The line (ray) is defined by 2 points P_0, P_1

$$s = P_0 + (P_1 - P_0)t \tag{4.2.23}$$

The triangle (environment) is defined by 3 points T_0, T_1, T_2

$$v = T_0 + (T_1 - T_0)m + (T_2 - T_0)n \tag{4.2.24}$$

The intersect is calculated by setting $s = v$

$$P_0 - T_0 = (P_0 - P_1)t + (T_1 - T_0)m + (T_2 - T_0)n \tag{4.2.25}$$

Other checks to determine if the line is parallel to the plane or if the intersection falls within the triangle are calculated in the software.

A method for improving the ray-shooting algorithm to increase the ray density was considered during early experiments, where the use of a computer's graphics card was used instead of the Central Processing Unit (CPU). Most graphics cards today have significant parallel processing power in order to generate the 3D environments common in games and applications. These cards can be reprogrammed to compute other algorithms and can be much more powerful than a CPU for certain calculations. The initial results were promising however a number of problems with compatibility between the toolboxes available and the graphics card used were found. This is certainly due to the immaturity of the field and should improve dramatically in the short term. Since this experiment was undertaken a number of developments have occurred already, making processing on a graphics card much more feasible.

4.3 Ionospheric propagation models

4.3.1 Introduction

As stated in Chapter 2, the ionosphere affects the GPS signal in a number of ways, but the most significant are the code group delay (combined with the carrier phase advance) and the scintillation. Scintillation is due to the ionosphere causing a diffraction grating effect which splits and merges different signal paths causing rapid shifts

in amplitude and phase. Group delay is caused by changes in refractive index, which is related to the electron density. At the time of development the reference models for the ionospheric delay and scintillation effects were International Reference Ionosphere model (IRI) and WBMOD respectively. This section describes the two models used.

4.3.2 International Reference Ionosphere

The IRI is an international project sponsored by the Committee on Space Research (COSPAR) and the International Union of Radio Science (URSI). The IRI Model is an empirical model based upon experimental measurements from various sources to determine the properties of the ionosphere. IRI can determine monthly averages of the electron density, total electron content (TEC), electron temperature, ion composition and vertical ion drift. For the GPS simulator only the Total Electron Content (TEC) is used to determine the signal propagation delay. Further information can be found in *Bilitza et al.* (1993); *Bilitza* (1997, 2001).

4.3.3 WBMOD

WBMOD is a radio-wave ionospheric scintillation model, developed at North West Research Associates, Inc. It was developed using measurements collected by the Defence Nuclear Agency (DNA) Wide Band satellite, which transmitted at frequencies from UHF to L-band. The WBMOD model provides estimates of amplitude and phase scintillation based on climatological model of the global distribution and behaviour of the ionospheric F-region plasma-density irregularities. The scintillation levels are calculated using the power-law phase screen propagation model. The phase screen model is based on the solution of the Fresnel-Kirchhoff diffraction approximation, which is applied over a screen of waves with Gaussian distributed phase. The Gaussian phase represents the irregularities in the ionosphere, in reality these irregularities occur through the ionosphere however this model approximate them to a single infinitely thin screen. The properties of the Gaussian distribution are determined by the climatological model. Applying the Fresnel-Kirchhoff integral to these waves determines the diffraction effects of the misaligned waves. WBMOD takes into account the solar cycle, season, geomagnetic disturbance levels, and local time. Further information on the model can be found in *Rino* (1979); *Secan et al.* (1995, 1997). There are a number of more accurate propagation models to this now, including the PWE method combined with phase screens or other random media models, however this method is a tried and tested method which provides acceptable results.

4.4 The Simulation

4.4.1 Signal Generation overview

- Simulation Software Development

- Basic Core
  - Satellite orbits
  - Earth model
  - Coordinate system
  - PRN codes
  - Timing
- Multi-path Model
  - Ray-Shooting Algorithm
  - Connect new rays to the previous ones
  - Apply delays and losses to signals
- Receiver DSP Model
  - Carrier Tracking loop models
    - Discriminator types
    - Bandwidth
  - Code Tracking loop models
    - Discriminator types
    - Bandwidth
  - Acquisition Algorithms
    - Standard
    - FFT based
- Signal Generator Model
  - Generate thermal noise
  - Generate receiver clock errors
  - Generate Signals for each path calculated by the ray-shooting algorithm
    - Calculate Doppler
    - Calculate delay offsets (path & ionosphere)
    - Apply PRN codes

        * Apply scintillation and noise
    - Antenna & RF stages model
        * FFT for filtering and group delay
        * Automatic gain controller
        * Sampling and Quantisation
- Ionosphere Model
    - Calculate group delays using
        * IRI
        * TEC values
    - Generate Scintillation using
        * WBMOD
        * S_4 & σ_ϕ values
        * Real data

The simulation uses the high precision satellite data files (SP3) that provide information about the orbital parameters of the satellites. The C/A code is generated using a software equivalent of the circuit in Figure 2.6. The navigation message is not used in the simulation however a repeating message can be input for testing bit error rates. The ray-shooting algorithm uses the satellite positions and receiver environment to determine the ray vectors and outputs the paths taken by any rays that reach the receiver. The new paths are related to previous paths based on the similarity of the paths taken, if the paths are dissimilar the new path is assumed to be a new signal. The path delay is calculated from the path distance and added to the delay from the ionosphere model (IRI), which determines the time offsets between the signals which reach the receiver. Clock errors are calculated based on the user inputs of the receiver's clock noise and drift. The received power is determined by the Friis free-space path loss equation with the addition of Fresnel coefficients to determine power losses due to reflections (4.2.18). The circularly polarised reflection coefficients are calculated by combining two linear polarised equations (*Hannah*, 2001). The gain and group delay of the receiver RF/IF chain can be specified as a function of frequency. Background noise power can be input as a single value which is typically calculated using:

$$N = kTB \tag{4.4.1}$$

where:

- k is the Boltzman constant.

- T is the temperature which, for a sky facing antenna with adequate ground attenuation is approximated as 100K.
- B is the bandwidth of the signal.

During initialisation, a signal is generated for each path determined by the ray-shooting algorithm. The signal is generated at an IF of 1.4 MHz sampled at 5 M samples per second, as generation at the transmitted frequency (1575 MHz) would require excessive processing and memory requirements for little benefit. The Doppler frequency for the carrier and the code is calculated from the change in path distance from the ray-shooting algorithm and interpolated to the sample rate. The amplitude and phase scintillation is interpolated to the sample rate and the background and receiver clock noise are generated. The carrier and code phases are calculated based on the Doppler and noise sources, as well as the initial offset due to propagation delay. The signal from each satellite is generated using the following equations:

$$S_t = (sin(\omega t + \Phi).C_{C/A}.D_{nav}.P_{recvd}) + A_{scint} + N_{backgnd} \tag{4.4.2}$$

$$\Phi = N_{clock} + D_{clock} + P_{scint} \tag{4.4.3}$$

where:

- S_t is the signal generated
- ω is the angular frequency of the carrier, in this case $2\pi \times 1.4 \times 10^6$ rad/s
- t is the time based on the sample rate, in this case 5×10^6
- Φ is phase offsets due to the receiver clock noise and phase scintillation
- $C_{C/A}$ is the C/A code value, 1 or -1
- D_{nav} is the navigation data, 1 or -1, but not typically used currently
- P_{recvd} is the signal power received by the receiver, based on path losses
- A_{scint} is the Amplitude scintillation interpolated to the sample rate
- $N_{backgnd}$ is the background or thermal noise generated from Gaussian and multiplied by kTB input value
- N_{clock} is the receiver clock noise generated from Gaussian and multiplied by user input value
- D_{clock} is the receiver clock drift generated from integrated Gaussian multiplied by user input value
- P_{scint} is the phase scintillation interpolated to the sample rate

The signals from each path are summed together creating the GPS L1 C/A code signal.

The receiver front end is modelled as part the the signal generator as saving the signal at this point would require a significant amount of storage (20 MB per second of signal). The gain and group delay of receiver front end is applied using the FFT. The AGC and quantisation is then applied creating a 2-bit 5 MHz signal (1.2 MB per second of signal). This is an ideal point to save the signal as the next step is the DSP stage of the receiver, which typically have the most varying designs, see Chapter 3. By doing this a comparison of different receiver algorithms can be made on the same data. Much of the simulator is very closely based on the real system, as the aim was to develop an investigative tool rather that a simulator that produces defined outputs. Many of the settings described here can be varied, but this may have an effect on the execution time.

4.4.2 Incorporation of the Ionosphere Refraction Model

The IRI model used provides vertical TEC, which is the integrated electron content directly above the position. To convert the vertical TEC into the slant TEC, which is the electron content between the receiver and satellite, the following approximation is used.

$$TEC_{SLANT} = \frac{TEC_{VERT}}{cosec(\frac{\pi}{2} - \theta_{SAT})} \tag{4.4.4}$$

where the θ_{SAT} is the satellite elevation.
The code group delay and carrier phase advance are equal in magnitude and opposite in polarity. They can be calculated from TEC value using the following equation.

$$\delta t = \frac{40.3}{cf^2} \int N.dl \tag{4.4.5}$$

where:

- δt is the ionospheric group delay or the carrier phase advance.
- c is the speed of light
- f is the frequency
- $\int N.dl$ is the TEC

These periods are added to the carrier phase and subtracted from the code phase of each signal in the signal generator.

4.4.3 Incorporation of the Scintillation Model

There are three possible ways to input scintillation into the simulator: Measured data directly from a receiver at 50 Hz ; S_4 and σ_ϕ values; the WBMOD model. S_4 and σ_ϕ are the one-minute standard deviations of amplitude and phase measurements respectively. They are used as a metric of the scintillation severity.

Measured data

The measured data must be detrended and normalised. The amplitude scintillation must have an average value of one and the phase must have an average value of zero for the simulator. The phase measurements are detrended using filtering based on the methods shown in *Dierendonck et al.* (1993) and missing values are interpolated using the spline method. This type of input is used to determine how a receiver reacts to specific scintillation fades from a real receiver.

S_4 and σ_ϕ

These values can be input as a single value, a single value for each satellite or a value for each satellite every minute. The values are converted into scintillation using the established noise distributions of Nakagami-m (*Nakagami*, 1960) for amplitude and Gaussian for phase. The noise sources provide outputs at a user defined bandwidth, which is typically 0.1-0.5 Hz (*Fu et al.*, 1999). These are then interpolated using the spline method to 50 Hz , then linearly interpolated to the receiver sample rate. This can be used when the full measured values are unavailable or for experimenting with combinations of amplitude and phase scintillation.

WBMOD implementation

The WBMOD model was converted from a command line input system to a MATLAB function for integration into the simulation. The model can provide a number of forms of output, however in the simulation it is limited to one. The user inputs a number of time percentages to form a statistical distribution of the maximum S_4 and σ_ϕ values. The S_4 and σ_ϕ values for a particular time are determined from the time distribution by random but with thresholds proportional to the percentage of time values. These are then converted to a 50 Hz scintillation using the same method as the S_4 and σ_ϕ input values. The WBMOD method is used for a general forecasting of the conditions.

Scintillation Analysis

The properties of the simulated scintillation have been compared to real receiver scintillation. Figure 4.2 shows the outputs for an amplitude scintillation case for the: real data (blue), simulated from S_4 (green); simulated from WBMOD (red). The results show that by setting the appropriate bandwidth in the scintillation simulator a good statistical match to real scintillation is possible. Unfortunately, as the bandwidth of real scintillation changes, no generic model can be determined. This method allows the analysis of combinations of amplitude scintillation, phase scintillation and bandwidth.

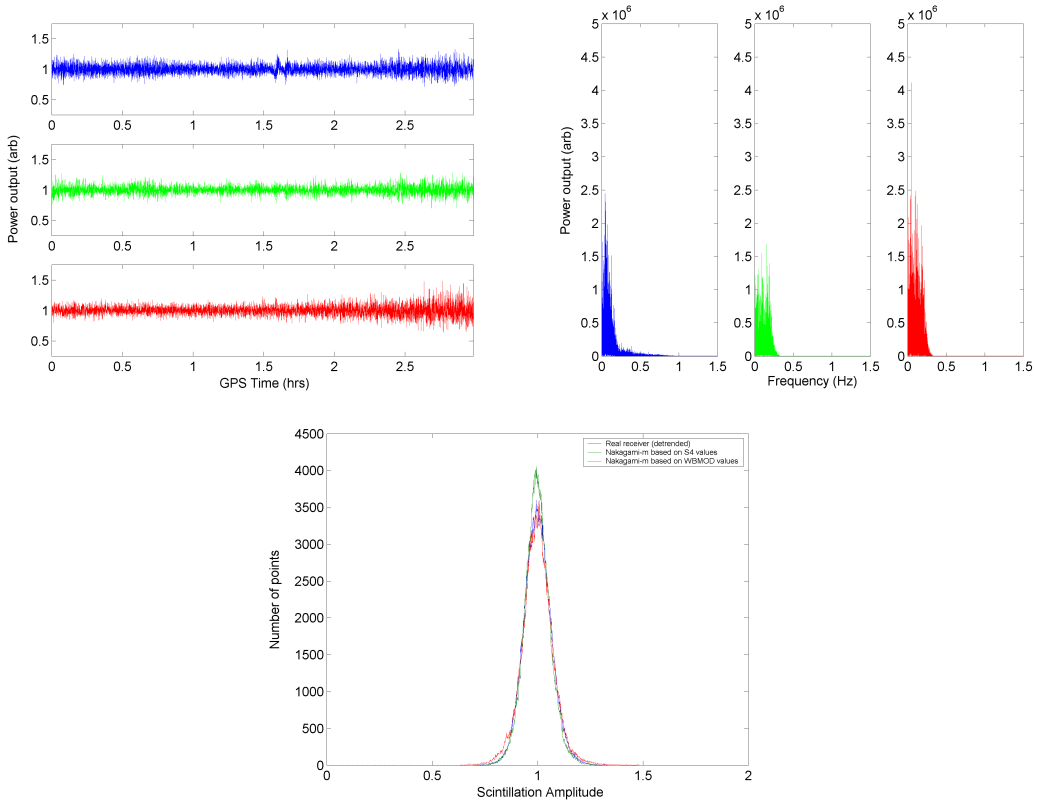


Figure 4.2: Shows the amplitude scintillation, distribution shape and the frequency spectrum for real, S_4 Nakagami-m and WBMOD Nakagami-m

4.4.4 The GPS DSP Receiver

Figure 4.3 shows a diagram of a typical GPS receiver which this software receiver is based on. The GPS receiver reads the data from the files created by the signal generator. The acquisition is done by the code accelerated FFT method described in Chapter 3. The code and carrier values from the acquisition are passed to the signal tracking part of the receiver. The signal tracking is based on a standard early-minus-late DLL for code-tracking, a FLL for carrier pull in and low C/N_0 tracking and a PLL for standard

carrier-tracking. The DLL is a carrier-aided second-order loop which has the option of being coherent or non-coherent. The PLL is a third order loop with either a standard Costas (multiplier), an arc-tangent or a decision-directed discriminator. The FLL is a second-order loop and uses a differenced form of the chosen PLL discriminator. The bandwidth of all the tracking loops can be adjusted using the input structure. The integration time for the loops is 1 ms but can also be adjusted. The outputs from the receiver are the I and Q power values, the C/N_0, the carrier phase error, the code-tracking frequency and the carrier-tracking frequency.

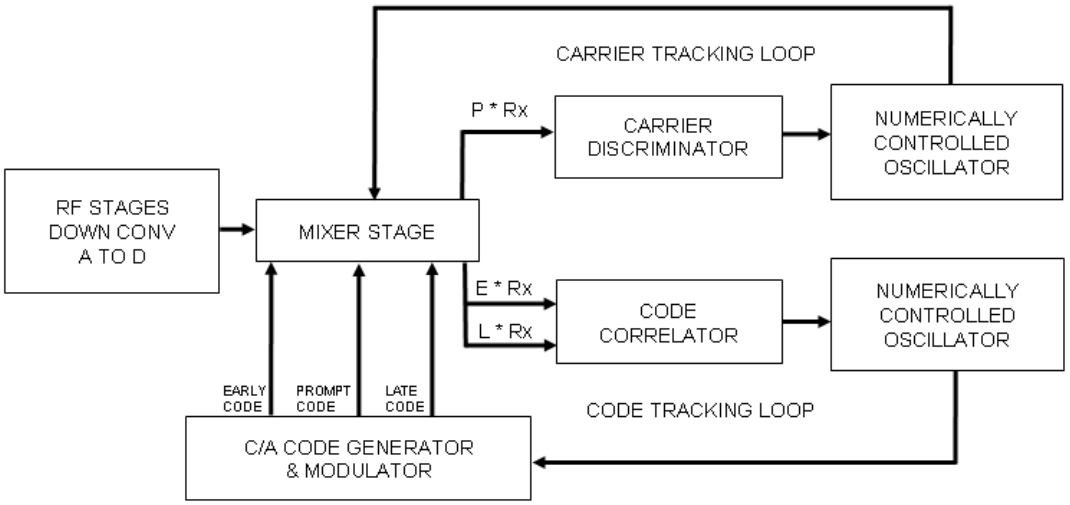


Figure 4.3: Diagram of a GPS receiver

4.5 Summary

This chapter has looked at how the simulation was developed. A number of radio propagation models have been described and evaluated, resulting in the chosen method for the simulator. The ray-shooting GO based model was shown to be the most suitable in this application. A description of the third-party ionospheric models and how they were integrated into the simulation has been given. The algorithm used by the simulator has been described, including the possible inputs and the effects modelled.

Chapter 5

Ground Multi-path Simulation

5.1 Introduction

This chapter shows comparisons between the outputs from a real receiver with those from the simulator under multipath conditions. The effectiveness of the simulator to model multi-path is evaluated, showing its abilities and limitations. The multi-path conditions were created by a reflective panel, mounted 30 cm east of the real receiver position. The real receiver is based on a roof top in the University of Bath campus so there exists a possibility of other multi-path sources. Figure 5.1 shows the position of the reflector to the receiver. Two different materials were used for the real reflector, both were 60 cm^2 panels, one of aluminium and the other a PVC mirror. The power and phase measurements from the real receiver are compared with those from the simulations.

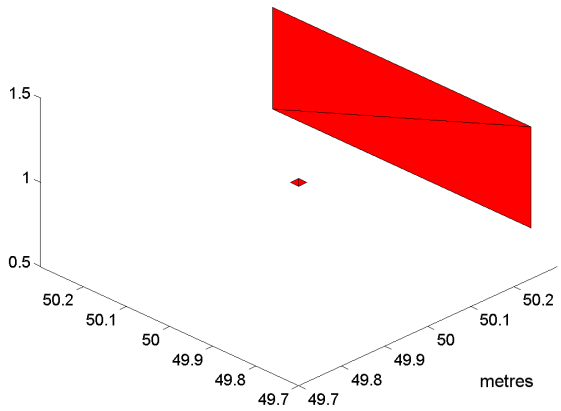


Figure 5.1: Diagram of the simulated receiver and the reflective panel.

5.2 Results and Analysis

5.2.1 Effect of the reflective panel on the power output

The real receiver used in this experiment is a GPS ionospheric scintillation monitor developed by Cornell University's Space Plasma Physics group (*Beach and Kintner*, 2001). This receiver can provide raw 50Hz power and integrated carrier phase measurements. No direct units can be determined for the power as the AGC and integration stages of the receiver remove any direct relation. There are methods for calculating C/N_0 proposed by Van Dierendonck in *Parkinson and Spiker* (1996) and covered in Chapter 4, however these require in-phase and quadrature values which are not available from the Cornell receiver. Therefore the simulator receiver is set up to match the parameters of the Cornell receiver (where possible) so the power outputs are comparable.

Figure 5.2 shows the power outputs from the real receiver on successive days from the 26th April 2006. Graphs (ii) and (iii) show the output when one of the reflective panels was installed, where as (i) and (iv) show the output when no panel was installed. A significant difference can be seen in the outputs when the reflective panel was installed. It is also noted than there is little difference between the outputs from the two different panels, the aluminium (ii) and the PVC mirror (iii). So despite the dielectric properties of the PVC mirror, it is the conductive mirror plating that is the significant factor in this case. The effect of the reflector is seen as a sinusoidal pattern in the received power between the times of 08:00 and 09:00, with a period of just over one hour. This pattern is caused by constructive and destructive interference between the carriers of the direct and reflected signals, where they move in and out of phase at the receiver antenna (see Figure 5.3).

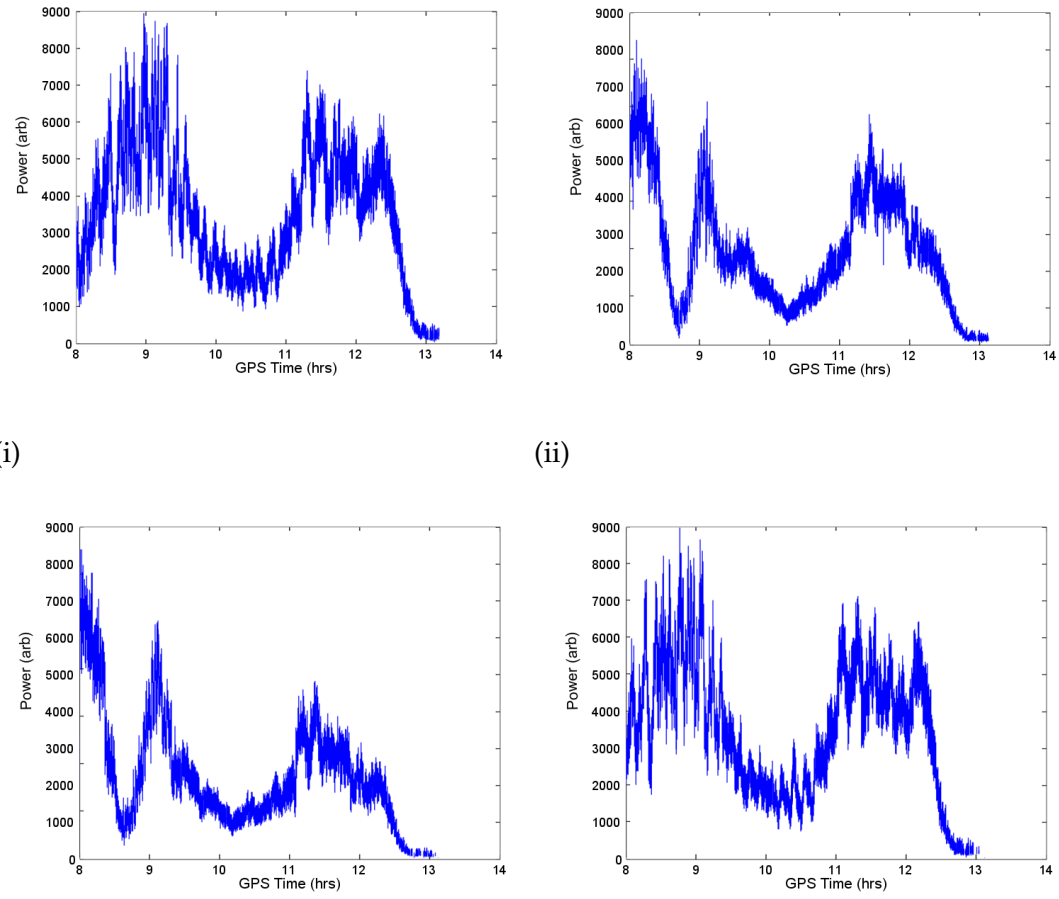


Figure 5.2: Power output for PRN11 from the real receiver on successive days starting on the 25 April 2006

A wide trough is seen in the results when the satellite is at a high elevation (GPS time of 10:30), rather than the expected peak as the satellite is at its closest to the receiver. The trough occurs whether the reflector is installed or not, ruling this out as the cause. The shape of the trough is similar to that of the interference caused by the reflective panel but of a longer period, suggesting that it is also caused by multi-path rather than an atmospheric effect. The longer period of this interference suggests the reflective surface is closer to the receiver antenna than the installed reflective panel. The only structure this close to the receiver is the weather-proof radome enclosure, shown in Figure 5.4. The simulation was used to model the reflective panel and then to investigate the effects of the radome.

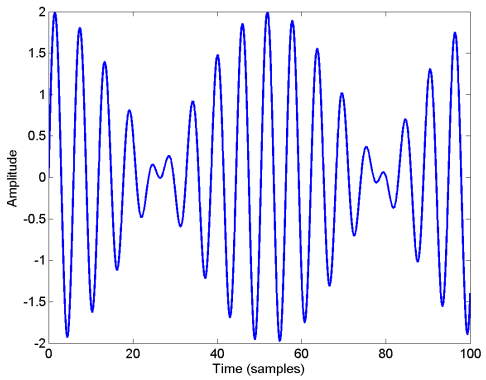


Figure 5.3: Plot showing two sinusoids of slightly differing frequencies summed. Constructive and Destructive interference shown

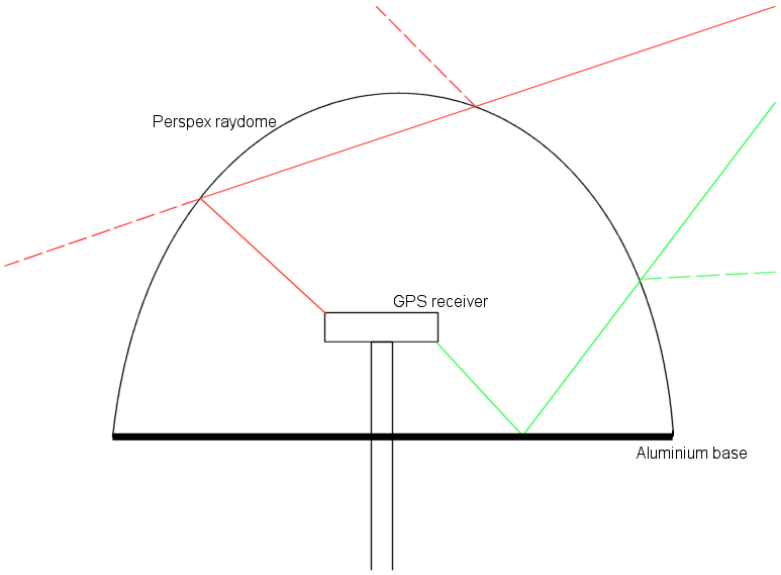
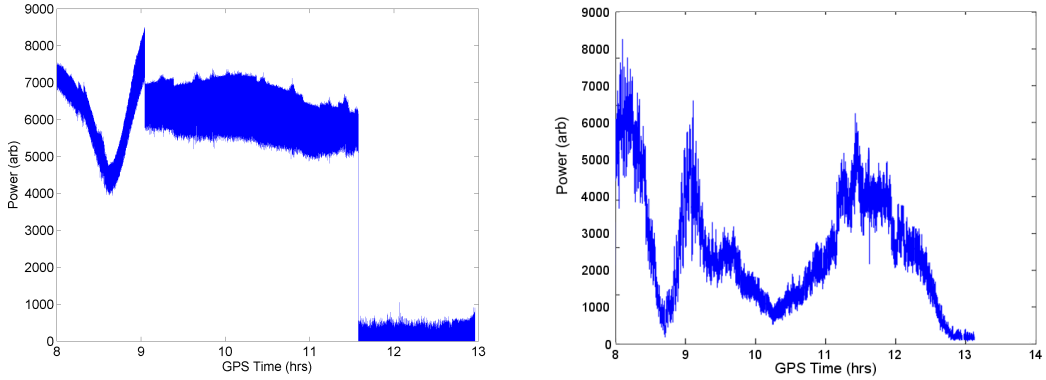


Figure 5.4: Diagram of the real receiver structure and possible ray paths

The simulation was used to attempt to replicate the multi-path interference, due to the reflective panel, seen in the real receiver's output. The simulation environment was set as shown in Figure 5.1 and the simulated receiver properties were set similar to the real receiver's properties where possible. No ionospheric effects were included in the simulation as these could be confused with the multi-path effects. The panel's reflective properties were modelled using Fresnel reflection coefficients and approximate values for the conductivity and permittivity obtained from (*Kraus and Fleisch*, 1999).

Figures 5.5 (i) and (ii), show a similar interference pattern between 08:00 and 09:00; the period and shape of the distortion is similar in both. The proportions show that

the simulation output is less affected by the multi-path than the real receiver. In the simulation the interference causes a 50% drop in output power where as a 75% drop is seen in the real receiver power output. This could be due to inaccuracies in the ray tracer based propagation model or differences in the receiver properties. However, the likely cause is a difference in the antenna's LHCP rejection ratio, used to attenuate singly reflected signals.



(i) (ii)

Figure 5.5: Plots showing the receiver power output from the: (i) simulation with the reflective panel, (ii) real receiver with the reflective panel

The simulation shows that the signal travels a secondary route via the reflective panel (shown in Figure 5.1), which causes the distance travelled to vary from the direct path. This causes the carrier signals from the two paths to merge after travelling different distances, causing the carriers to move in and out of phase. The simulation also shows a sharp drop in power and a loss of lock when the satellite is at low elevation at approximately 1130 UT. This is caused by the simulated ray being blocked by the reflective panel, unlike the real case where it can diffract around it, causing a more gradual drop.

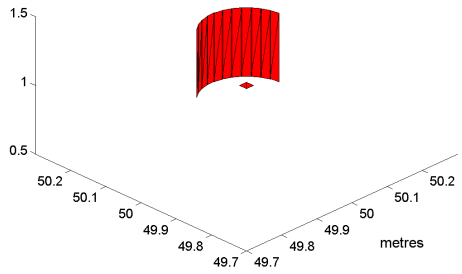
5.2.2 Effect of the radome on the power output

As discussed earlier in the chapter, the trough in the power output of the real receiver is likely caused by part of the weather proof radome. This theory is based on the period of the interference pattern, if the secondary ray path is only slightly longer than the direct path then it can only vary slowly with satellite elevation. The radome consists of a perspex container and an aluminium base, either could cause reflections near the receiver. Due to its dielectric properties, the perspex container can reflect and refract the signal with varying ratios depending on the ray's angle of incidence. Refraction of the signal by the perspex container is unlikely to cause enough of a delay to inter-

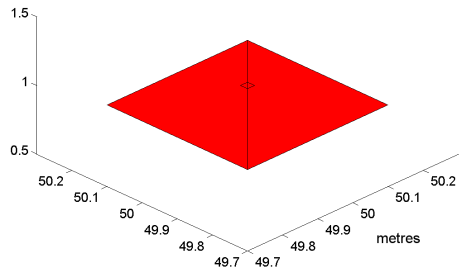
fere with the carrier, as the path would be almost identical to the direct path, so only reflections were considered. The base of the radome is made of aluminium providing a constant reflection coefficient for all angles. Many GPS antennas are shielded at the base and sometimes the sides (choke ring), rejecting any ray arriving up to a certain elevation. The antenna used was not of this type, therefore would not reject signals arriving from the radome base. To determine which part of the radome causes the trough, simulations were executed using both environments shown in Figure 5.6, (i) the perspex container, and (ii) the base of the radome.

Figure 5.6 (iii) shows the receiver power output from a simulation of the perspex enclosure (Figure (i)). The effect of the perspex is small and only evident at the start and end of the satellite pass. The radome is conical shaped with a rounded top, it is approximately half a metre tall and has the reflective properties shown in Figure 5.7 for a GPS signal. Due to the slope of the conical enclosure only rays arriving from low elevations can reflect back to the receiver antenna. Outputs were taken from the real receiver with and without the perspex radome, however no noticeable differences were seen. The simulation shows some sharp changes in the power output when the interference occurs, this is a known problem with ray-based algorithms where discrete changes are seen in the output. This problem is reduced if the ray density can be increased, however this increases the computational requirement substantially.

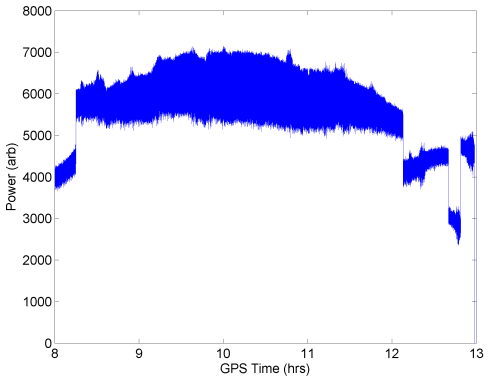
Figure 5.6 (iv) shows the receiver power output from a simulation including the radome base (ii). The output shows a large trough with a sharp peak in the centre as the satellite reaches its maximum elevation. The period of the trough is similar to that of the real receiver, however no central sharp peak is seen in the real receiver output. This can be explained by the antenna model used in the simulation; the size of the plane used to represent the antenna is related to the ray spacing of the ray tracer, which is larger than the real antenna. This larger plane causes the secondary ray, reflected from the base, to be blocked for a short period when the satellite is at high elevation. By blocking the multi-path, the interference is removed and the output power increases to its normal value. In addition to this reason, the real electromagnetic field can diffract around the antenna maintaining a secondary path. Other than the central peak, the simulation output shows a good similarity with the real receiver output, proving that the trough is caused by the metallic base of the radome.



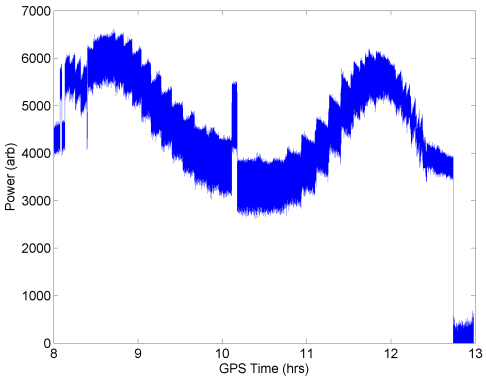
(i)



(ii)



(iii)



(iv)

Figure 5.6: Plot showing the receiver environments (i) the glass radome enclosure, (ii) the radome base, followed by the receiver power output from the simulation from each environment respectively.

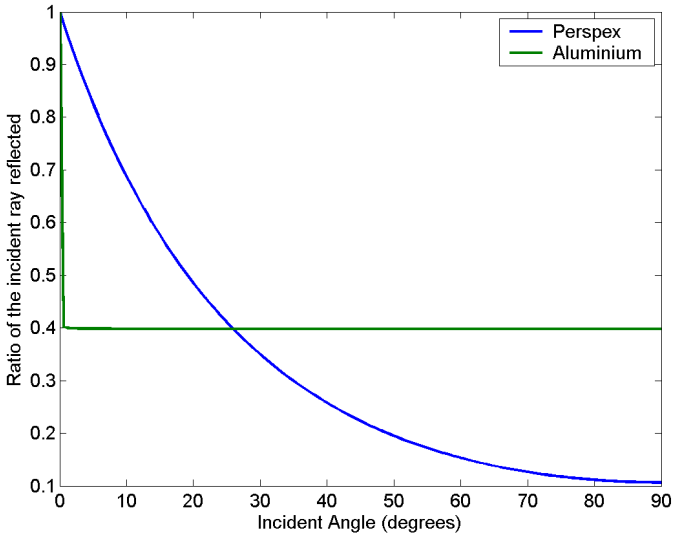
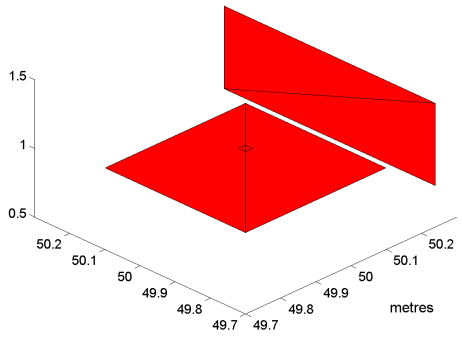


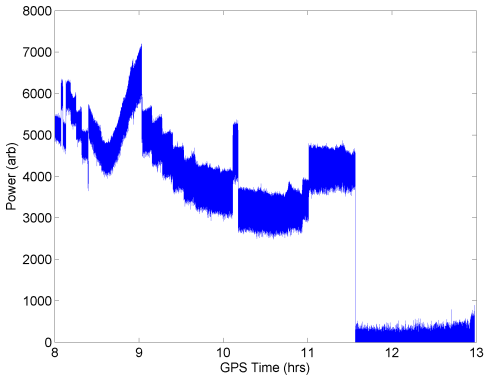
Figure 5.7: Graph showing the perspex radome and the aluminium base RHCP reflection coefficients with varying ray incidence angle

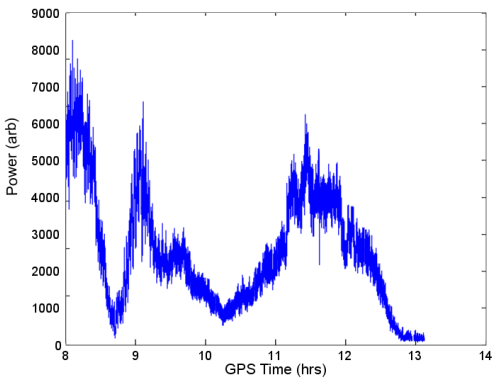
5.2.3 Effect of the reflective panel and the radome on the power output

The final part of this case study is to simulate the full environment to reproduce the real receiver's output. Figure 5.8 shows the environment and power outputs for the simulated and the real receiver. The main features are present in both results and show a good match, however there are a few discrepancies. The first is the rapid fluctuations in the simulated output at the beginning of the measurements. These are caused by approximating the radome base in the simulation as a square rather than the real circular base. As a square is not of equal diameter at all azimuth angles, unlike a circle, it causes the effective reflector length to change as the satellite moves in azimuth. This is not a problem for most cases, only when a ray is reflecting off the edge of the reflector and the satellite is moving significantly in azimuth. This constant changing of base length and elevation causes the fluctuations in the simulation. This is not a problem with the simulation and was done to save processing time. Another discrepancy is the intensity of the multi-path effect which is evident in all the simulations and was discussed in the results and the cause was determined to be antenna differences. Finally, the sharp cut-off at approximately 11:30 is due to the direct ray being blocked by the mounted reflector also mentioned previously. Figure 5.9 shows some of the outputs from the ray-shooting algorithm in each of the environments investigated.



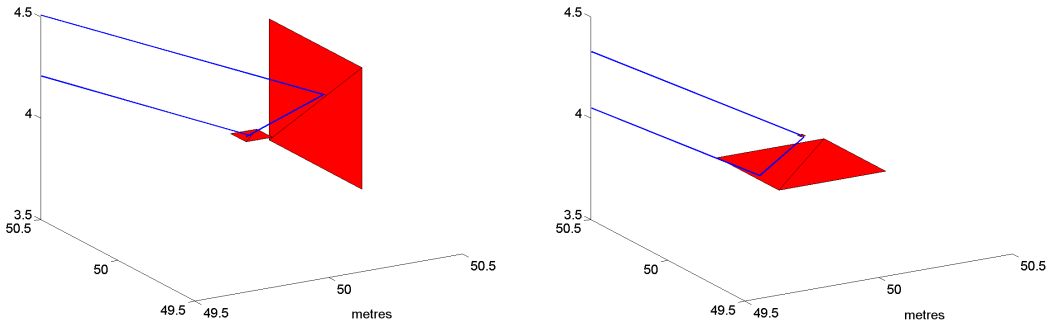
(i)



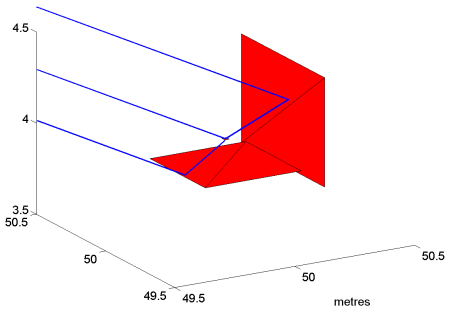


(ii) (iii)

Figure 5.8: Plots showing the effect of the base and mounted reflector shown in (i) on the simulated receiver (ii) and the real receiver (iii).



(i) (ii)



(iii)

Figure 5.9: Plots showing the routes taken by the rays to the receiver in the different environments (i) mounted reflector. (ii) radome base. (iii) mounted reflector and radome base.

5.2.4 Effect of the reflective panel on the phase output

Figure 5.10 shows the phase output from the real receiver on successive days from the 25th April 2006. Graphs (ii) and (iii) show the output when one of the reflective panels was installed, where as (i) and (iv) show the output without any panel. There is a dip in the phase variance at 08:45 corresponding with the trough in the power output (Figure 5.8), which is attributed to the reflective panel. There is also a shallow dip in the phase variance corresponding to the wider trough in the power output attributed to the radome base. No cycle slips are seen in the results, an abrupt change of over 0.5 cycles would signify this. The multi-path has little effect on the phase discriminator as the changes are slow and are not strong enough to completely remove the signal.

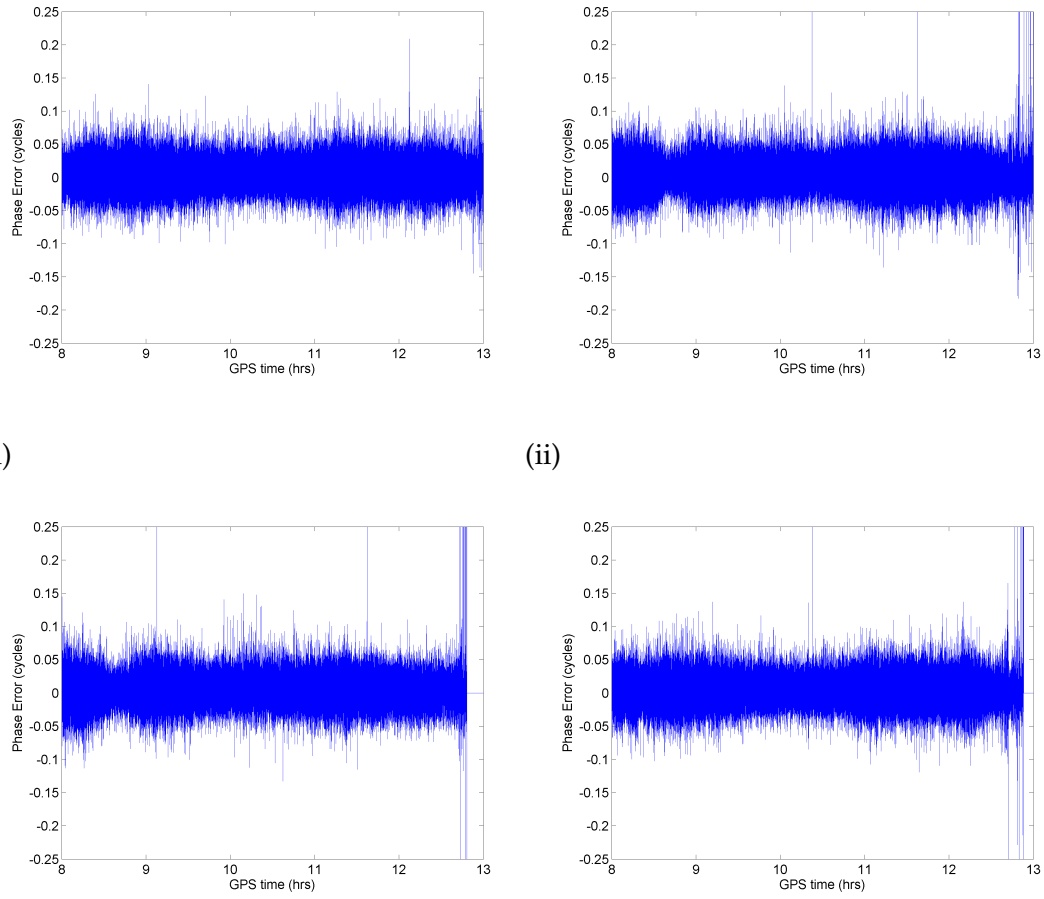
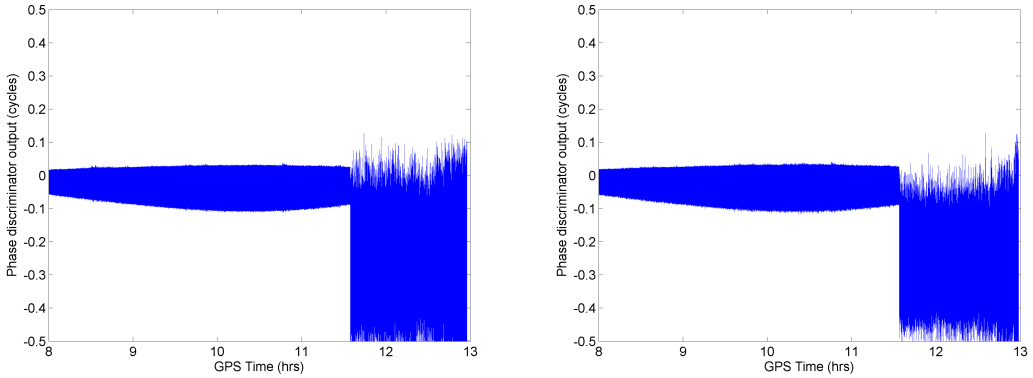


Figure 5.10: Phase outputs on successive days from the real receiver i)Without reflector ii)With reflector iii)With reflector iv)Without reflector

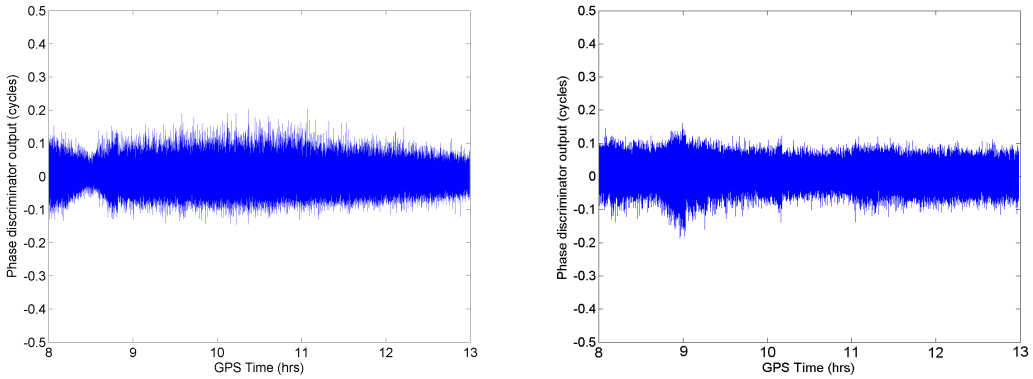
Figure 5.11 shows the phase discriminator output of the receiver from the simulation. It can be seen that the opposite effect is seen to the real receiver, where the phase variance increases slightly when the power reduces. This difference was traced to the type of carrier discriminator used, where the real receiver uses a Costas multiplier discriminator, the simulation uses an arc-tangent type. More information on how the PLL operates can be found in Chapter 2 (The GPS Receiver). The Costas multiplier produces an output directly related to the power of the input signal, which directly effects the output variance. The arc-tangent discriminator calculates the angle between the in-phase and quadrature values and therefore is independent of the input power, however as the power decreases the background noise increases causing the variance to increase. To verify this is the cause, a simulation using the standard Costas based discriminator was ran. Figure 5.12 shows the phase output from the Costas simulation for both the mounted reflector and combined reflector environments. It shows a very similar output to the real receiver proving that this was the cause. This simulation shows the ability

of the simulator to analyse receiver characteristics, as well as different environments.



(i) (ii)

Figure 5.11: Phase outputs from the simulated receiver using an arc-tangent discriminator with i) mounted reflector, ii) mounted reflector and radome base.



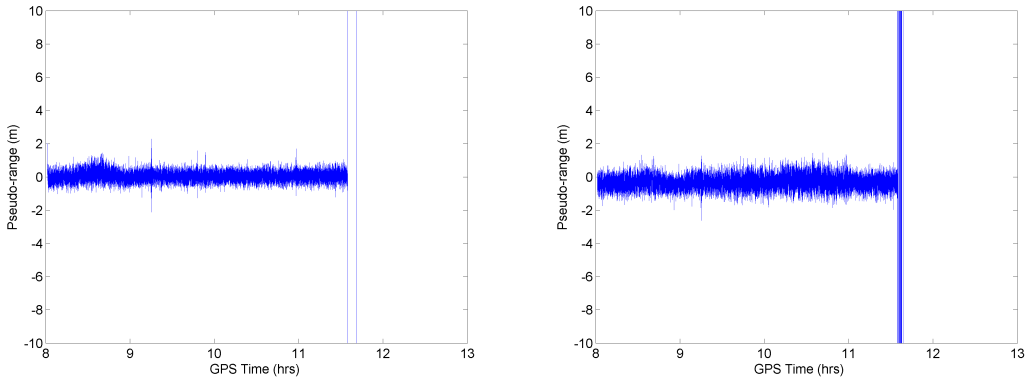
(i) (ii)

Figure 5.12: Phase outputs from the simulated receiver using a Costas discriminator with i) mounted reflector ii) mounted reflector and radome base.

5.2.5 Effect of the reflective panel on the positioning

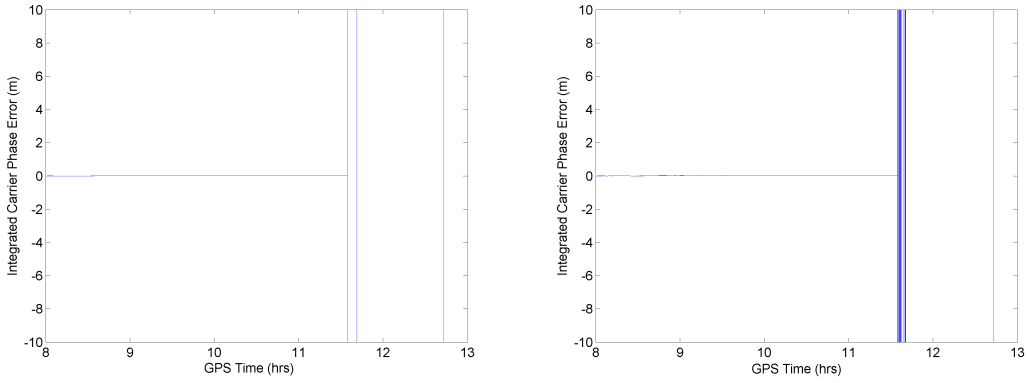
The previous section has shown the effect multi-path has on the receiver's ability to track the code and carrier of the signal. This section shows how the multi-path affects the receiver's positioning accuracy using either the pseudorange or the integrated carrier phase. Both the code pseudorange and the integrated carrier phase are calculated for the multi-path and non-multi-path environments then differenced to determine any errors.

The Figures 5.13 and 5.14 show very little error from the multi-path environment. The integrated carrier errors are very low in both cases. The pseudorange peak error is approximately 0.5 m for just the reflector and 1 m for the combined environment. The greater error of the pseudorange in the combined environment is due to the destructive interference increasing the background noise power. This does not occur to the integrated carrier phase as it is tracking on a much finer scale, 19 cm wavelength compared to 293 m .



(i) (ii)

Figure 5.13: Pseudorange error from the simulated receiver with i) mounted reflector ii) mounted reflector and radome base.



(i) (ii)

Figure 5.14: Integrated carrier phase error from the simulated receiver with i) mounted reflector ii) mounted reflector and radome base.

5.3 Conclusion

The simulated results show a good match to the real receiver results and it has been possible to replicate the effects of the surrounding local environment. In the real receiver results, some of the smaller scale variations with periods of less than 2 minutes, shown in Figure 5.2 (i)&(iv) were not considered in this experiment as it would require modelling a detailed replica of Bath university's campus, which is currently impractical due to the limited processing power. This possible limitation to the simulation is not considered final and improvements to the ray-shooting algorithm should improve the results further and enable more complex environments to be modelled. The experiment, described in Chapter 6, determines the cause of the small scale fluctuations. Another difference between the simulated and real results was the magnitude of the carrier interference from the reflective panel, which was discussed in the results and attributed to differences between the antenna polarisation rejection properties.

The experiment has shown that carrier interference from multi-path can significantly affect the received power. Carrier interference is seen from both the base and mounted reflector, however no cycle slips or loss of lock occurred throughout the experiments. This suggests a single ray from a nearby reflector cannot cause any significant positioning errors in a receiver, as the reflected ray is always a fraction of the direct signal. In an urban environment (conductive reflectors) this fraction is determined mainly by the polarisation rejection ratio, where typical antennas have at least 3 dB rejection for a singly reflected ray. However, from the effect of a single ray on the power, as few as two rays could cause a loss of lock. It can be seen in the figures from the real and simulated scenarios, that in a static environment short range multi-path can easily be distinguished from other errors in the GPS by its sinusoid shape. It has also shown that the sinusoidal distortion of the signal has a period related to the distance of the reflector and the satellite elevation. This periodic feature could be key in identifying multi-path errors from others using the receiver outputs.

Multi-path can also affect the modulated code of the signal; however this was not seen in any of these experiments. To see this effect, there must be a much longer difference in paths between the direct and multi-path rays. This is due to the much slower clock rate of the code, where one chip has a wavelength of 293 m (carrier wavelength is 0.19 m). There are many different DLL discriminators which vary in their susceptibility to multi-path (see Chapter 3). In the typical 'Early-minus-late' discriminator a difference of quarter to half a wavelength is required to cause a significant error, shown in Figure 5.15 ('1 chip spacing C/A code'), which translates to a path difference of approximately 100 m to 150 m . One of the less susceptible discriminators, the narrow correlator is also shown in Figure 5.15 ('0.1 chip spacing C/A code').

These experiments have verified the ability of the simulator to model the effect of

nearby multi-path on a GPS receiver. Although there are still improvements that could be made, the experiments have shown that the simulator is a valuable tool for analysing receivers as well as receiver sites. The next area to investigate is the periodic feature of the interference and the effects of longer range multi-path.

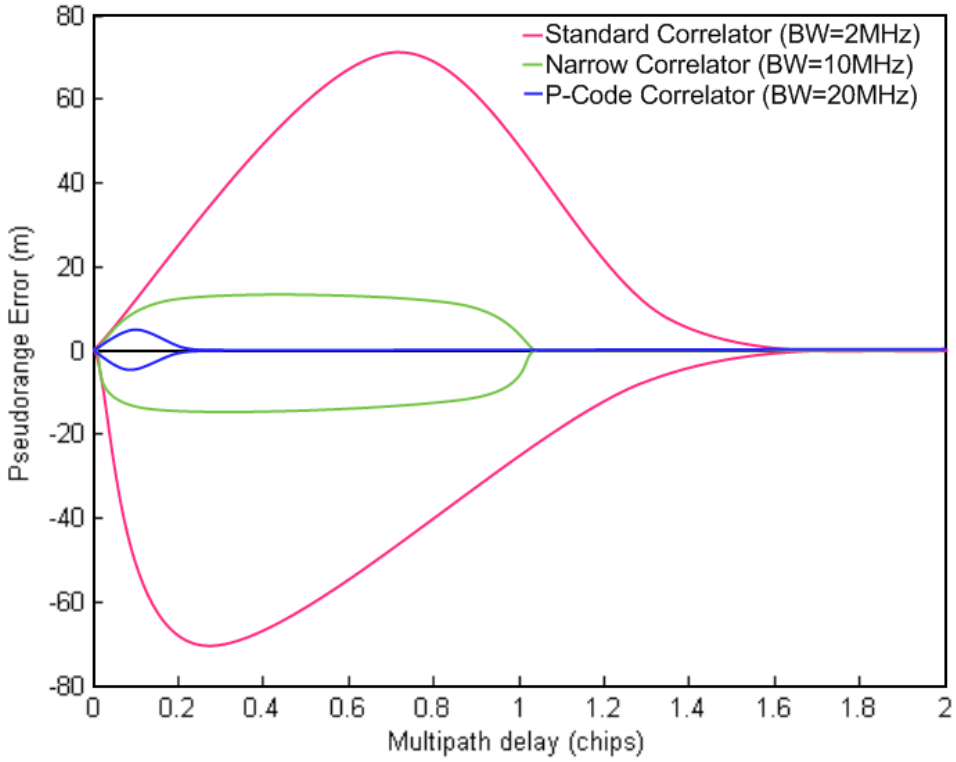


Figure 5.15: Shows the DLL Error against the delay difference of the direct ray and the multi-path ray (half direct power).

5.4 Summary

This chapter described an experiment where a real receiver was placed near a reflective surface to determine its effect on the receiver outputs. The GPS simulator developed was given a similar reflector in its modelled environment and the results were compared. The real receiver and simulated results were discussed and a strong similarity was shown. A further investigation of an unusual interference pattern in the real receivers output was completed, which provided an explanation of the cause. Finally, a discussion of the effects on the PLL and DLL was made.

Chapter 6

Analysis of the Multi-path Signature

6.1 Introduction

In the previous chapter only near-by multi-path reflections were considered, which cause low frequency fade patterns. This chapter looks at larger and more complex multi-path environments and determine how these effect the receiver output. The receiver outputs are analysed to determine a defining characteristic for multi-path, that can be used to identify it from other errors in the system. A method of identifying multi-path from the receiver signal power is presented and tested on simulated and real data.

6.2 Multi-path Simulation Experiment

6.2.1 Introduction

This experiment investigates the effects of varying the distance of the reflector from the receiver. As long-distance multi-path is difficult to set-up accurately on a real receiver, the simulator is used to investigate the issues. The experiment is based on an environment consisting of the receiver and a large conductive wall which is moved to provide different reflection positions.

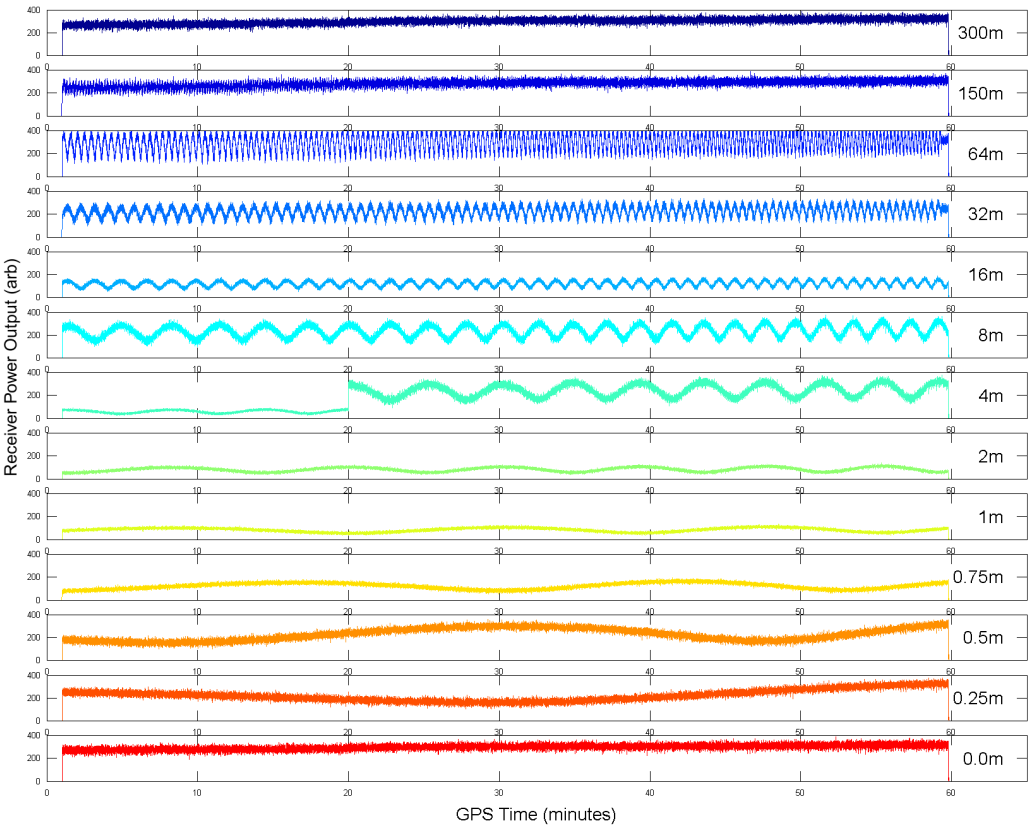
6.2.2 Results

Figure 6.1 (i), shows the power output from the simulation for the reflector at different distances from the receiver. It can be seen that as the distance between the reflector and receiver increases the period of the fade pattern reduces. This sinusoidal inter-

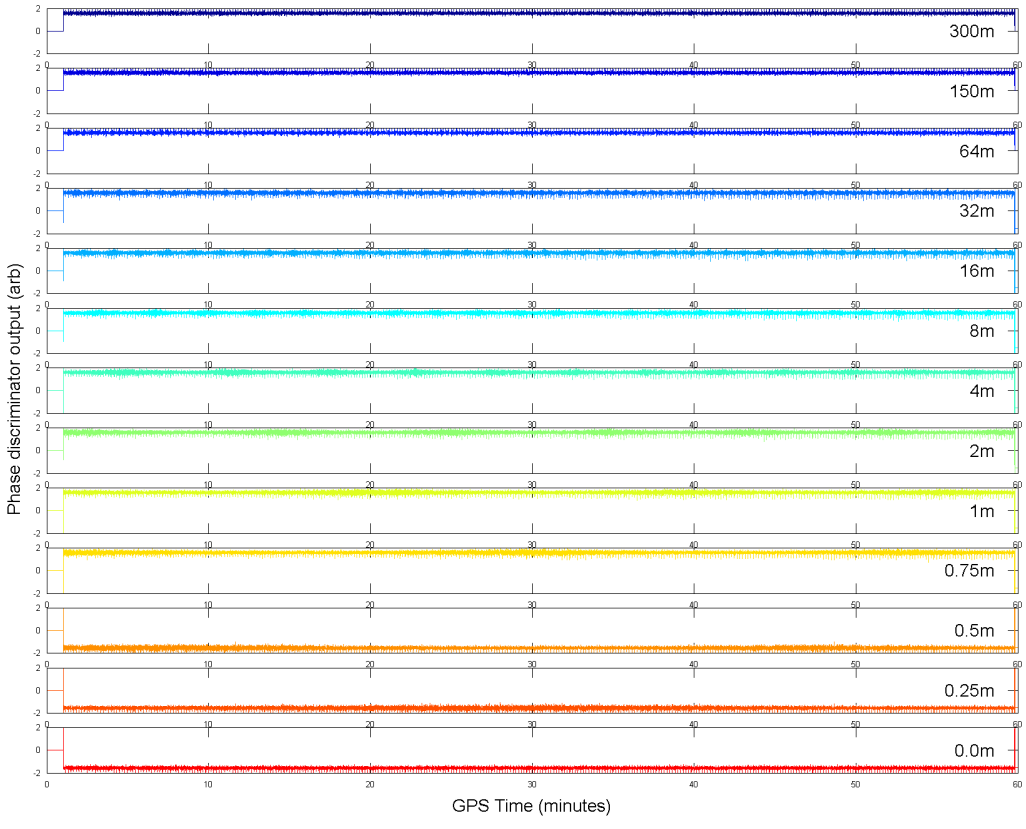
ference pattern is due to the carrier of the direct and indirect ray shifting in and out of phase. This relationship continues until the distance reaches 150 m where the interference reduces. This is because the code of the reflected signal is far enough out of alignment with the direct signal so that the DLL can reject it. A single code chip length is 293.2 m and the reflected ray from the 150 m reflector travels up to 300 m further than the direct ray, depending on satellite elevation. This is one chip cycle difference, which is shown in Figure 5.15 (1 chip spacing) to have less of an effect on the DLL.

Figure 6.1(ii) shows the phase error output from the PLL for the same reflector distances. This is the output from the carrier discriminator and has stable lock points at $-\frac{\pi}{2}$ and $\frac{\pi}{2}$ radians. If the output changes between stable points during tracking it signifies a cycle slip; constant changes signify a loss of lock. The results show no cycle slips or any evidence of the variations seen in the power output. This is confirmed by the Integrated Carrier Phase (ICP) (also known as the accumulated delta range) error, shown in Figure 6.2(i), where little or no error is shown. Unlike the carrier tracker the code tracker suffers from significant errors attributed to the carrier interference shown in the pseudorange error from Figure 6.2(ii). The ICP and pseudorange error are calculated by subtracting the range measurements of a clear environment from those of the multi-path environment. The oscillating characteristics in the pseudorange error suggests that the code discriminator is dependent on the carrier power as well as the code offset. To investigate this effect a simpler simulation of the DLL error slope was used to determine the effect of carrier interference on the code discriminator.

Figures 6.3(i)-(v) show the effect carrier interference has on the early-minus-late discriminator error slopes for different multi-path code offsets. Figure (i) shows the output when the direct and multi-path codes are in-line and the multi-path carrier phase is varied to show its effect on the discriminator. It can be seen that the amplitude of the error slope changes, however the shape and stable point remain the same. In Figure (ii) where the multipath code is out of alignment by half a chip (150 m difference), a significant range in the stable point can be seen, as the carrier phase is varied. This effect initially increases then reduces as the code delay difference is increased (Figures (iii) and (iv)). Figure (v) shows the complete separation of the direct and multi-path error slopes, removing any effect of the multi-path on the DLL.This shows that the carrier interference can significantly effect the DLL when the direct and multi-path codes are out of line by up to 1.5 chips. This code range implies that reflections from a distance greater than 300 m are unlikely to have any effect on the DLLs tracking ability.

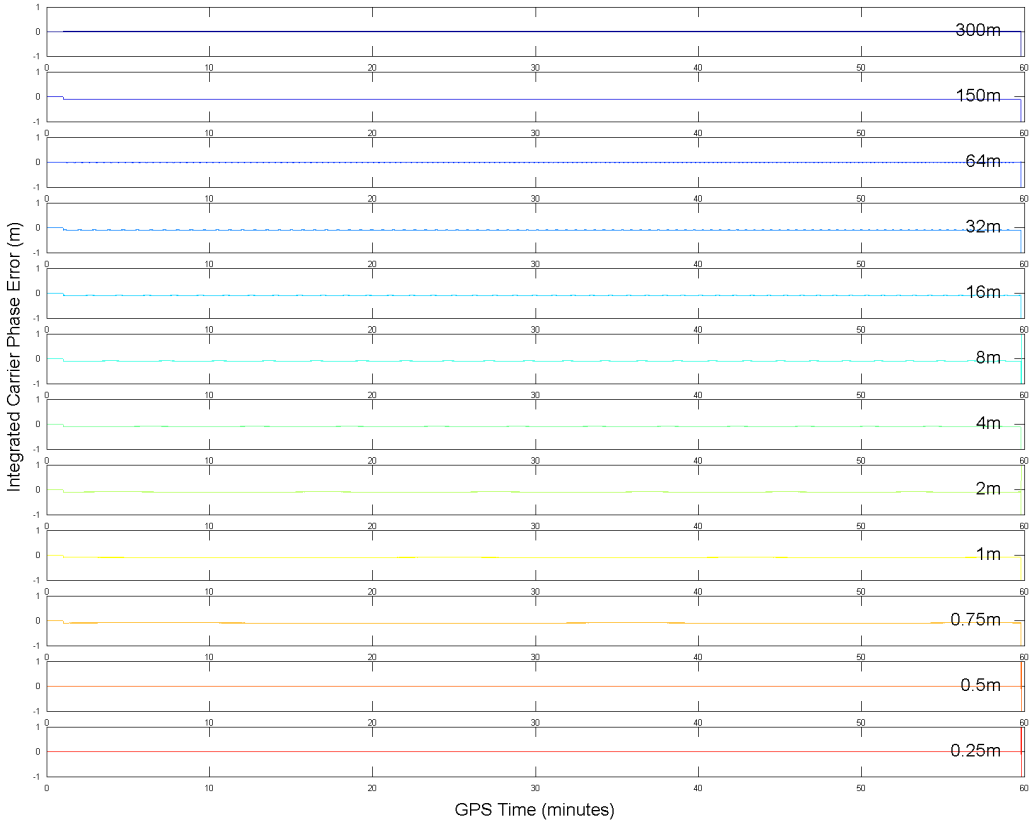


(i)

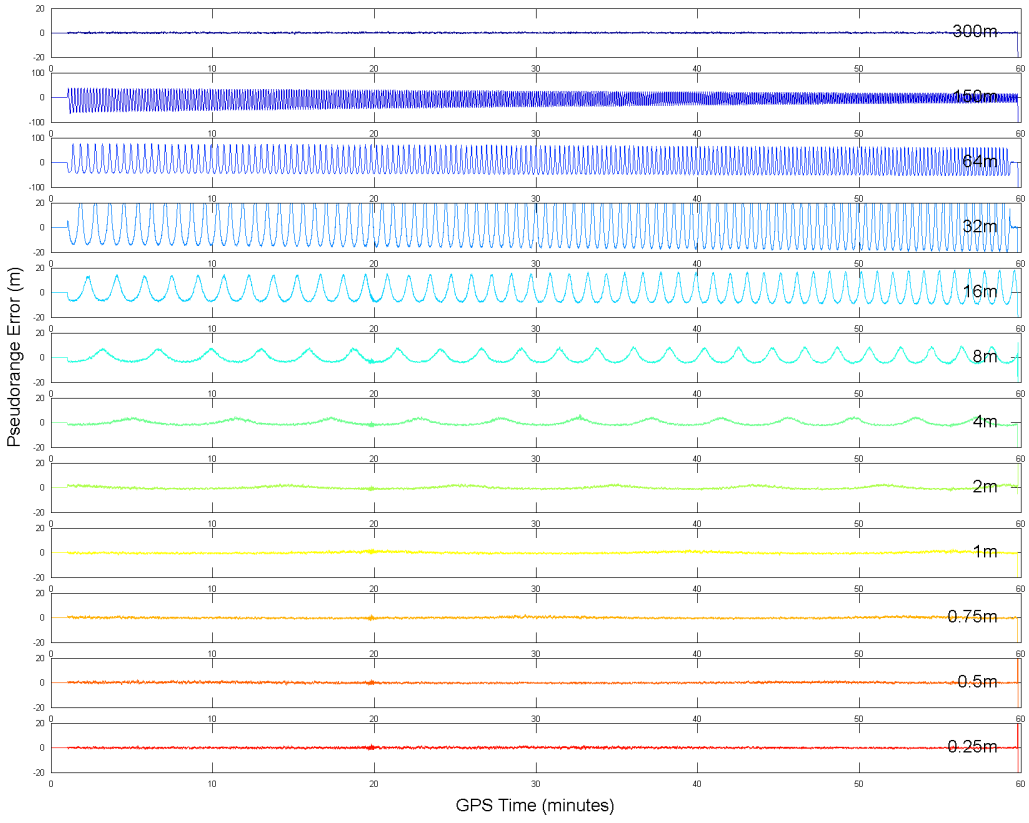


(ii)

Figure 6.1: Figures show the simulated receiver outputs, (i)Power (ii)Phase, for a reflector at different positions

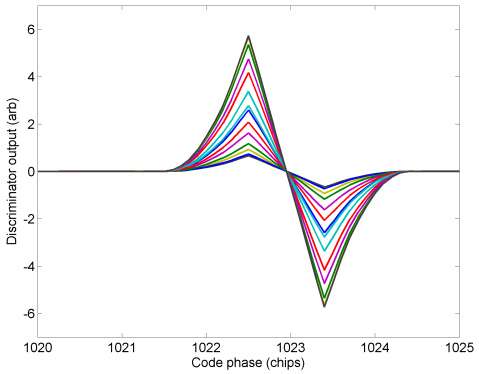


(i)

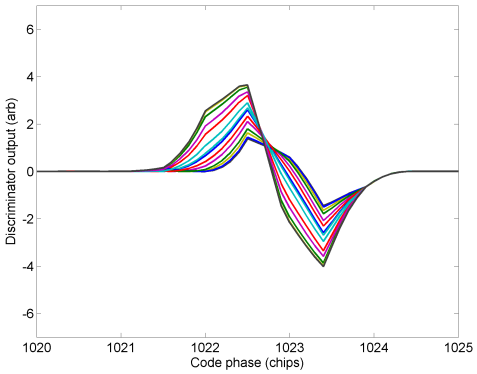


(ii)

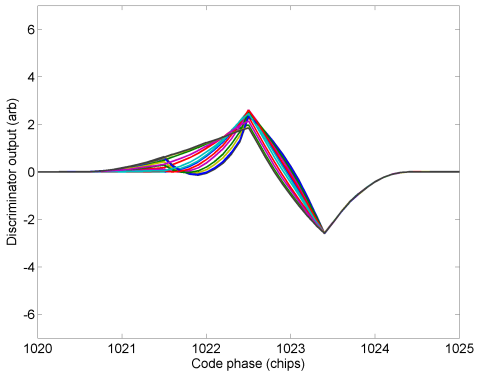
Figure 6.2: Shows the integrated carrier range error and pseudorange error for various reflector positions



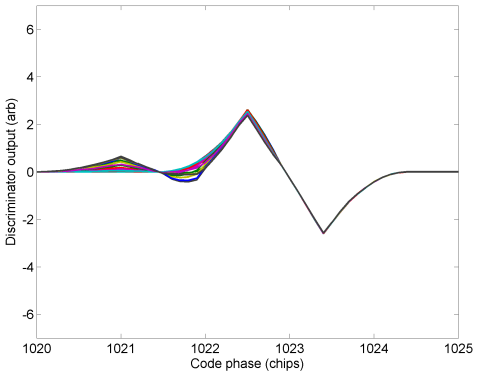
(i)



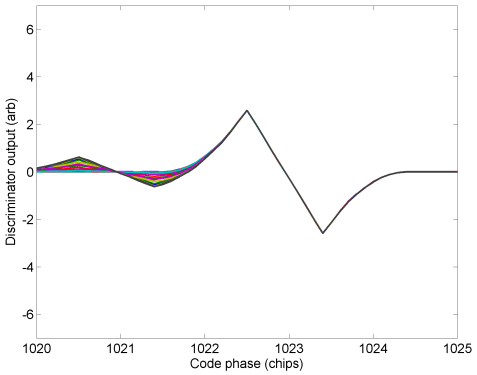
(ii)



(iii)



(iv)



(v)

Figure 6.3: Shows the DLL discriminator output for a direct and half power multi-path signal at different code delays and carrier phases- (i) 0 chips (ii) 0.5 chips (iii) 1 chip (iv) 1.5 chip (v) 2 chips

6.2.3 Analysis

This experiment has shown that there is a definite relationship between the period of the interference pattern and the distance of the reflector from the receiver. The relationship is based on the trigonometrical identities shown in Equation 6.2.4 and Figure 6.4. This makes the assumption that the direct and multipath rays travel in parallel from the satellite, which is valid for small angles between the multi-path and direct ray.

$$r_2 = \frac{x}{\cos\theta} \tag{6.2.1}$$

$$r_1 = \frac{r_2}{\cos 2\theta} \tag{6.2.2}$$

$$r_1 + r_2 = 2x\cos\theta \tag{6.2.3}$$

$$\tag{6.2.4}$$

$$a_1 = \frac{h}{\sin\theta} \tag{6.2.5}$$

$$a_2 = \frac{a_1}{\cos 2\theta} \tag{6.2.6}$$

$$a_1 - a_2 = 2h\sin\theta \tag{6.2.7}$$

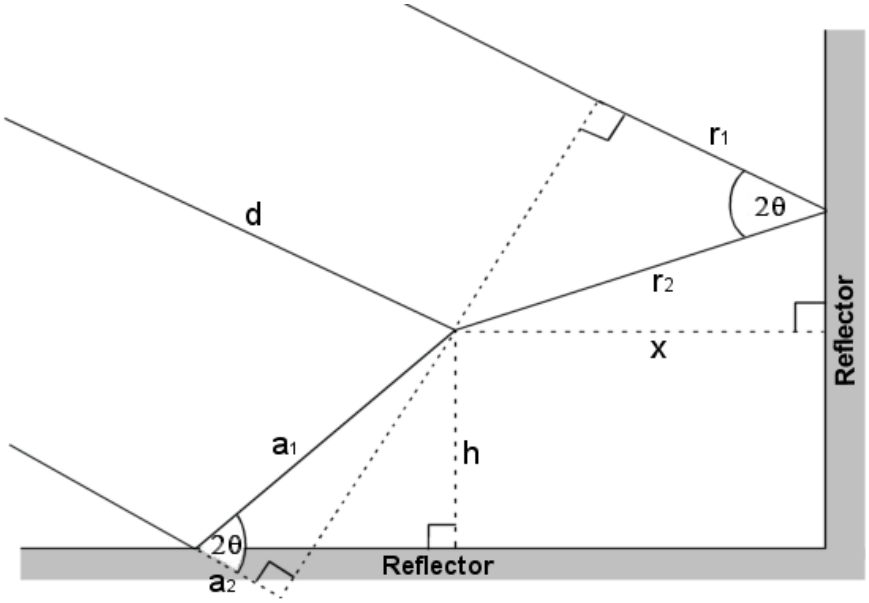


Figure 6.4: Diagram showing the 2-D trigonometry of multi-path

This shows that the difference between the direct and multi-path distance varies from

zero to twice the perpendicular reflector distance (x,h) as the satellite elevation (θ) changes from 0 ° to 90 ° . When the difference is equal to a multiple of the carrier wavelength (19 cm), constructive interference occurs, while an additional half wavelength offset causes destructive interference. In these experiments multi-path is easily identified by the oscillating changes in the signal power, however a method is needed to identify it in more complex situations where multiple rays reflect off multiple objects before reaching the receiver. The oscillating pattern suggests that multi-path maybe more obvious in the frequency domain, however the frequency of the oscillations change with elevation making it less viable. If the effect of the satellite elevation on the oscillations was removed the frequency would remain the same, which would be easily identified in the frequency domain. For the vertical wall case, interpolating the results to be a function of the cosine of the elevation (Equation 6.2.4), then transferring it to the frequency domain should make any multi-path evident. Figure 6.5 show the power output before and after the interpolation in the time and frequency domain, where the frequency axis is labelled as periods due to the very low frequencies. It can be seen that the multi-path is easily identified in the frequency domain after the interpolation and from this an estimation of the reflector distance can be made.

Unfortunately this method only works for reflections on vertical walls or horizontal floors in the case of the sine equation, but it can be adapted to different reflector orientations. By adapting the Discrete Fourier Transform (DFT) basis to a form that changes the frequency according to the above relationship, a transform that specifically analyses multi-path characteristics can be created. The transform can search the reflector distances and elevations to determine whether any multi-path occurs and the likely source of it. The basis of the DFT is changed from that shown in Equation 6.2.8 into that shown in 6.2.9, which oscillates in relation to the satellite elevation, θ. The transform varies the distance (x or h) and rotates the reflector (ϕ) to search the signal for multi-path signatures.

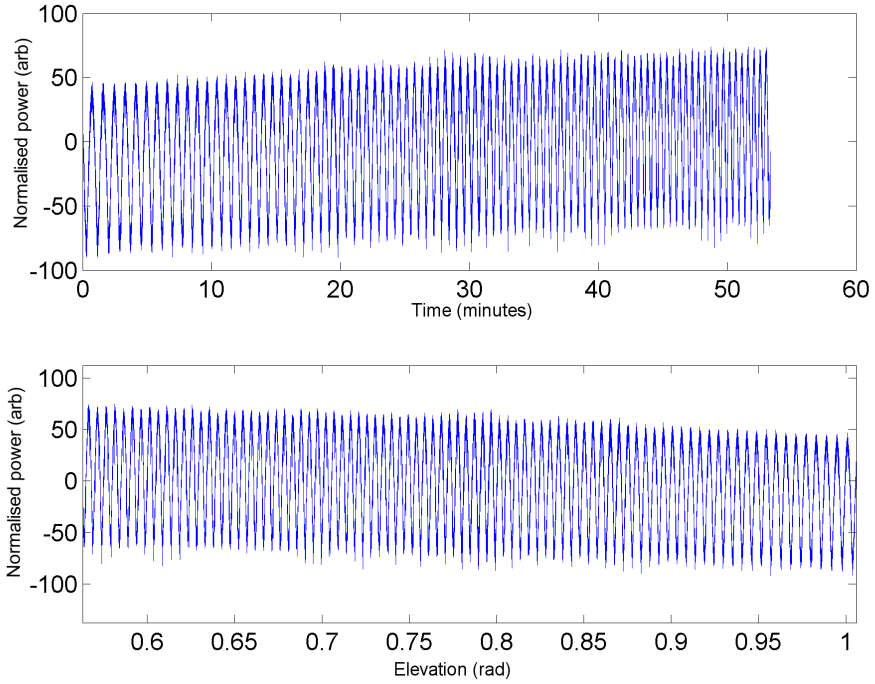
$$X_k = \sum_{t=0}^{T-1} x_t e^{\frac{-2\pi i}{T} kn} \tag{6.2.8}$$

$$X_{[d,\phi]} = \sum_{t=0}^{T-1} x_t e^{\frac{-2\pi i 2d \sin(\theta_t + \phi)}{\lambda}} \tag{6.2.9}$$

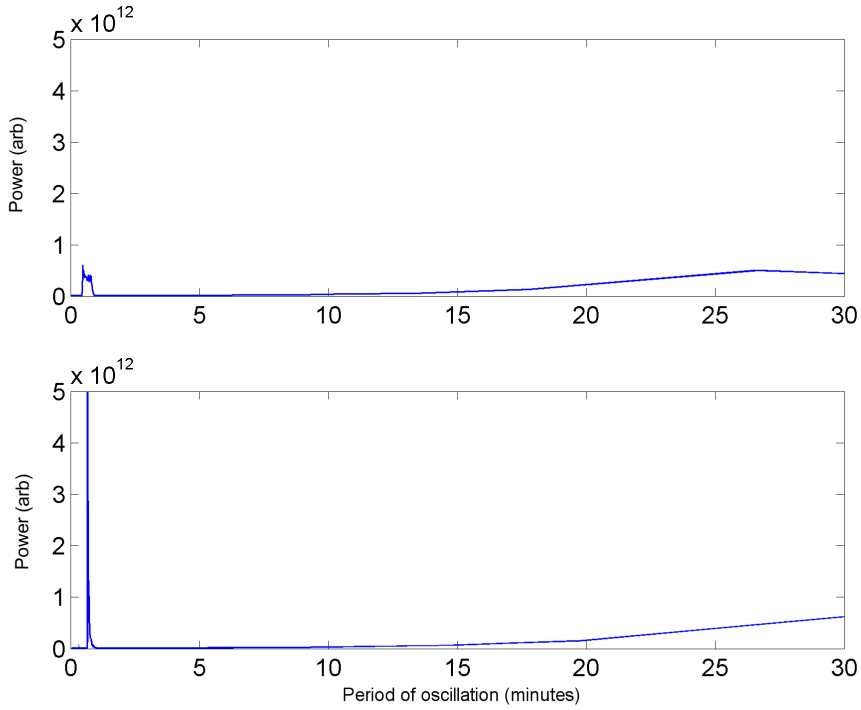
Where:

- k is the frequency index, t is the time index
- x_t is the signal in the time domain
- $X_{[d,\phi]}$ is the transformed signal
- d is the perpendicular distance to the reflector as shown in Figure 6.4
- θ_t is the elevation of the satellite at each sample of the signal

- ϕ is the elevation offset determined by the reflector orientation
- λ is the carrier wavelength (converts distance into phase)



(i)



(ii)

Figure 6.5: Shows the receiver output power in the time (i) and frequency (ii) domain before (top) and after interpolation (bottom) for the 32m reflector

Figure 6.6 shows the results of the Multi-path Analysis Transform (MAT) on the receiver power output for the vertical reflector at distances of 8 m and 32 m . A peak is shown at 9 m for the 8 m case and at 38 m for the 32 m case; the reflector angle is shown as 85 degrees in both cases. These are slightly different to the values expected, however they are correct but must be interpreted differently. This is because the relationship shown in 6.2.4 is only in the 2D plane, in a 3D environment this plane may shift in orientation making the analysis more difficult. In this case, the plane is 14 ° from the vertical causing the reflector elevation angle and distance covered to be slightly different from the vertical case. Despite the interpretation difficulties in determining where the multi-path came from, the transform is able to clearly identify multi-path signatures in these cases. The MAT is used to investigate more complex scenarios in the next section.

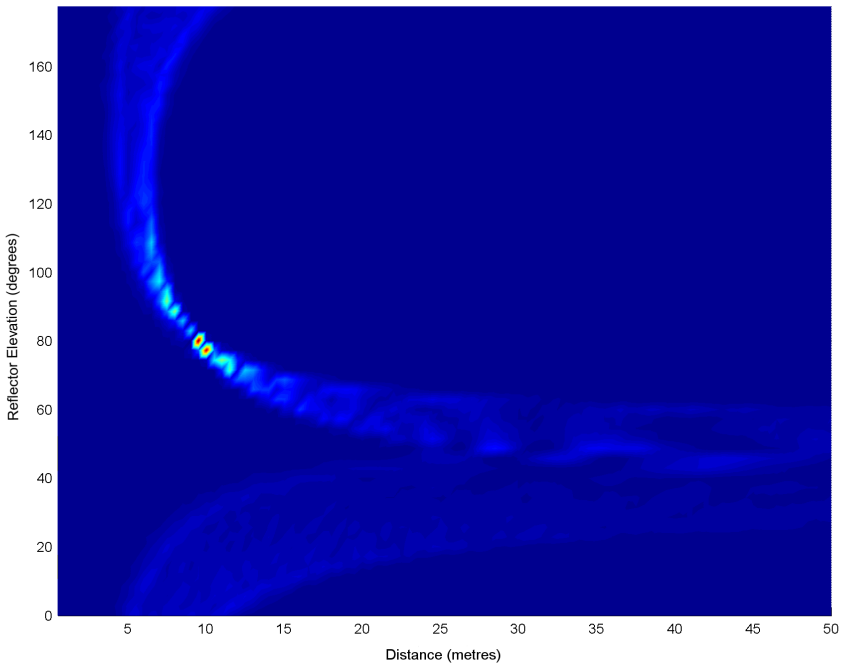
6.2.4 Conclusion

The experiment has shown that the period of the interference pattern is related to the distance between the reflection and the receiver. It has also shown a method to identify and analyse multi-path signatures in the power output of a receiver. Although there are some issues with interpreting the outputs in its current version, further development should reduce the problem.

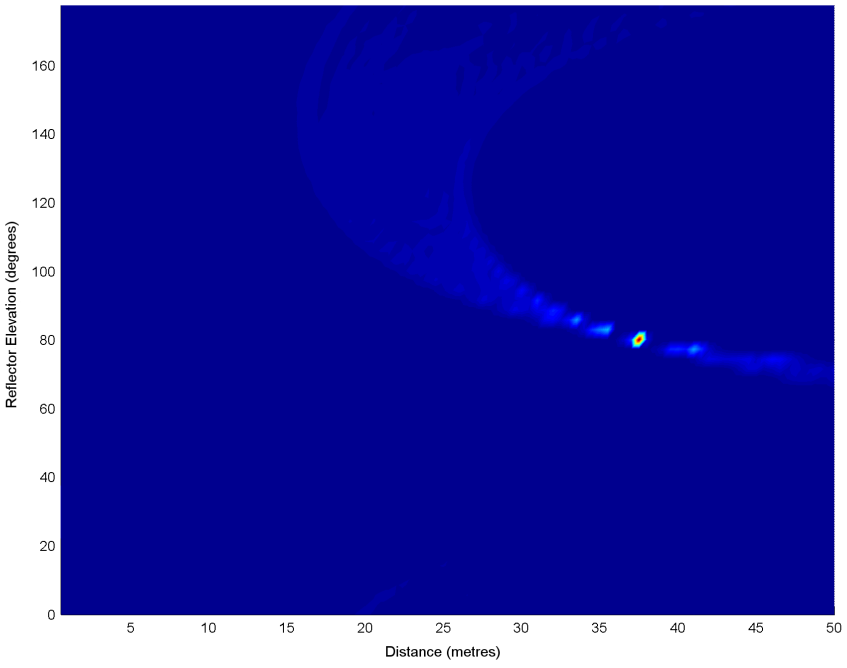
In this experiment carrier interference has had little effect on the PLL however, the DLL can show a significant error, depending on the alignment of the direct and multi-path codes. The results shown in 6.2(ii) show an increasing error as the wall reflector moves from 0.25 m to 64 m , then at 150 m (0.5 to 1 chip delay) the error reduces and is completely removed by 300 m (1 to 2 chip delay). This corresponds to the DLL error for multi-path shown in Figure 5.15, for a 1 chip spacing DLL, where the error starts to reduce at 240 m (0.8 chips) and is completely removed by 400 m (1.4 chips). Therefore distortion of the code is only evident when the reflection is less than 300 m (1 to 2 chips delay) away from the receiver. This is true for most cases, however if carrier interference or other errors combine with the long range multi-path, it is possible for the DLL to lose lock on the direct signal and begin tracking the multi-path signal. It should also be noted that these values are for the 1 chip spacing early-minus-late DLL and other discriminators may behave differently.

Multi-path has been shown to affect both the carrier and code of the signal, where the carrier interference produces the oscillating changes in power, and the code interference causes the oscillating errors to affect the DLL tracking. It is also shown that the carrier phase and integrated carrier phase have not been effected by the multi-path, confirming that carrier-tracking is a definite advantage in a GPS receiver. This chapter has so far analysed only single multi-path rays and has found that although they

have a significant effect on the receiver, in most cases they are not strong enough to cause large positioning errors. The next section investigates more complex scenes with multiple ray paths and analyse how these affect the receiver.



(i)



(ii)

Figure 6.6: Shows the transform outputs for both reflection types for a (i) 16 m reflector (ii) 32 m reflector

6.3 Multi-path Simulation Scenes

6.3.1 Introduction

So far, this chapter has investigated the relationship between the reflector distance and fade pattern period. A method has been derived to identify multi-path (the Multi-path Analysis Transform (MAT) method) using the receiver power output and has been tested on some simple multi-path environments. This section investigates more complex multi-path environments, where multiple reflections from multiple surfaces occur. As mentioned previously the ray-shooting algorithm is currently unable to handle very detailed environments, so some approximations have been made. For example urban buildings are modelled as cubes and rural cliffs are modelled using a triangular mesh. Each environment is shown as well as the receiver outputs of power, phase, integrated carrier phase and pseudorange. Any points in the receiver output that warrant investigation (marked by a red line) are analysed using the ray-shooting output, which shows all the rays that reach the receiver. The multipath analysis transform (MAT) developed in the previous section is also used to determine its effectiveness in more realistic environments.

6.3.2 Results and Analysis

Flat ground environment: Receiver without choke ring antenna. Satellite PRN 3

This simulation uses a receiver without a choke ring, or any type of low or negative elevation rejection, which is placed 2 m above a conductive ground medium. From the power graph in Figure 6.7 an oscillating pattern is seen, which changes in period as the satellite elevation increases. The phase discriminator output from Figure 6.7 shows a short loss of lock at 14 minutes, followed by some cycle slips at 23 minutes. These errors are caused solely by the multi-path causing destructive interference and dropping the signal power below the tracking threshold. This is unlikely to occur in a real situation as a single reflection would not be powerful enough, however it demonstrates the effect of severe carrier interference. The receiver recovers rapidly from the loss of lock without reverting to reacquisition, as it is able to switch to a wider bandwidth FLL in difficult conditions. The ICP and pseudorange show small errors even at very low power which may be attributed to the FLL.

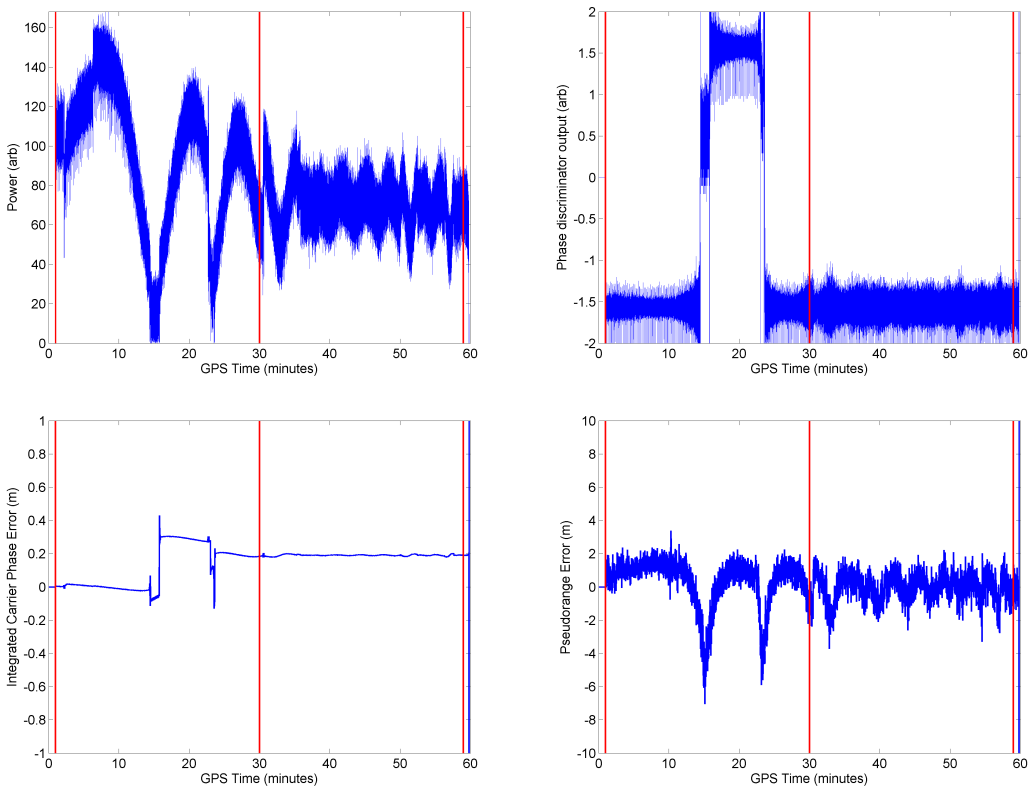


Figure 6.7: Receiver results from flat ground environment for satellite PRN 3.

Figure 6.8 shows the environment and the direct and multi-path rays for the start (0 minutes), middle (30 minutes) and end (60 minutes) of the measurements. The ray-shooting algorithm shows two indirect rays reaching the receiver from very similar reflection points, which is maintained intermittently through the simulation. This is a known problem with ray-shooting, where multiple rays from effectively the same reflection point reach the receiver and cause intermittent jumps and falls in power. This can usually be removed by careful design of the simulated environment, however this case has been retained as it has a significant effect on the receiver.

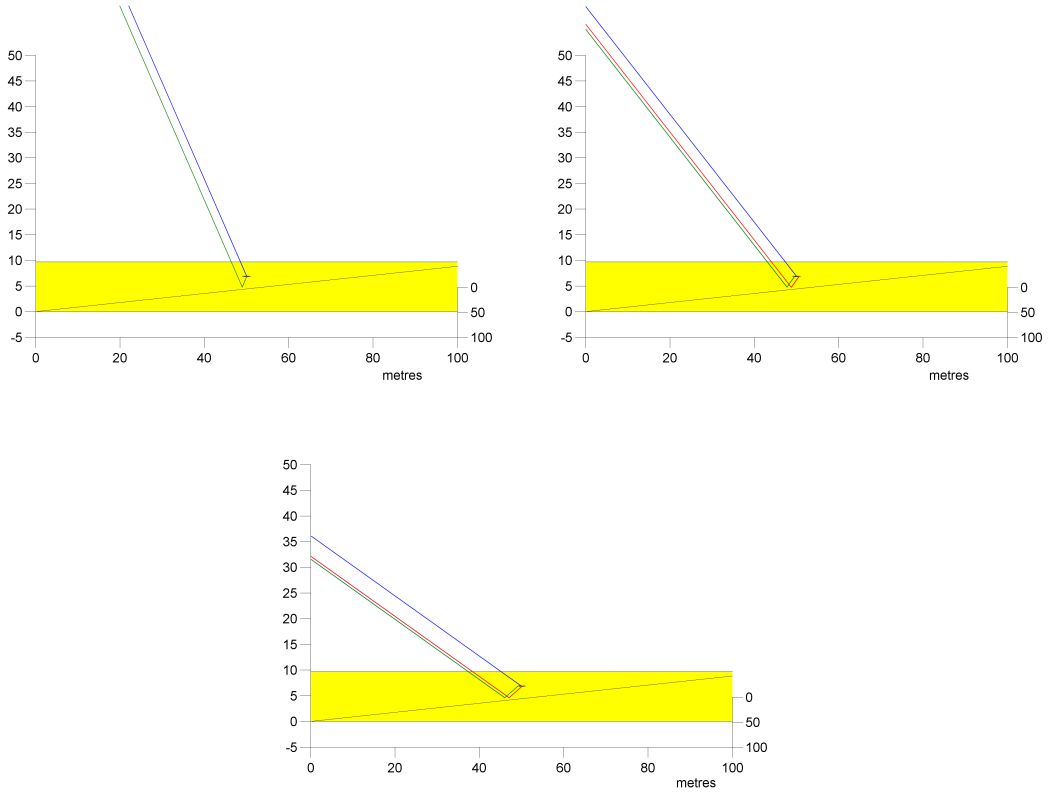


Figure 6.8: Environment and ray paths at 0, 30 and 60 minutes.

Figure 6.9 shows the MAT output for this case. The output shows a wide peak which varies from 2 m to 10 m at 0 ° to 30 ° , where the correct value should be 2 m at 0 ° . The wide peak may be attributed to either the two rays interfering with each other or the low number of oscillations shown in the output. This problem may be evident in all nearby reflections, however nearby reflections can easily be identified in the power output. In this case there are no interpretation issues with the 2D plane differing from the assumed vertical plane in multi-path analysis as the environment only has a horizontal reflector.

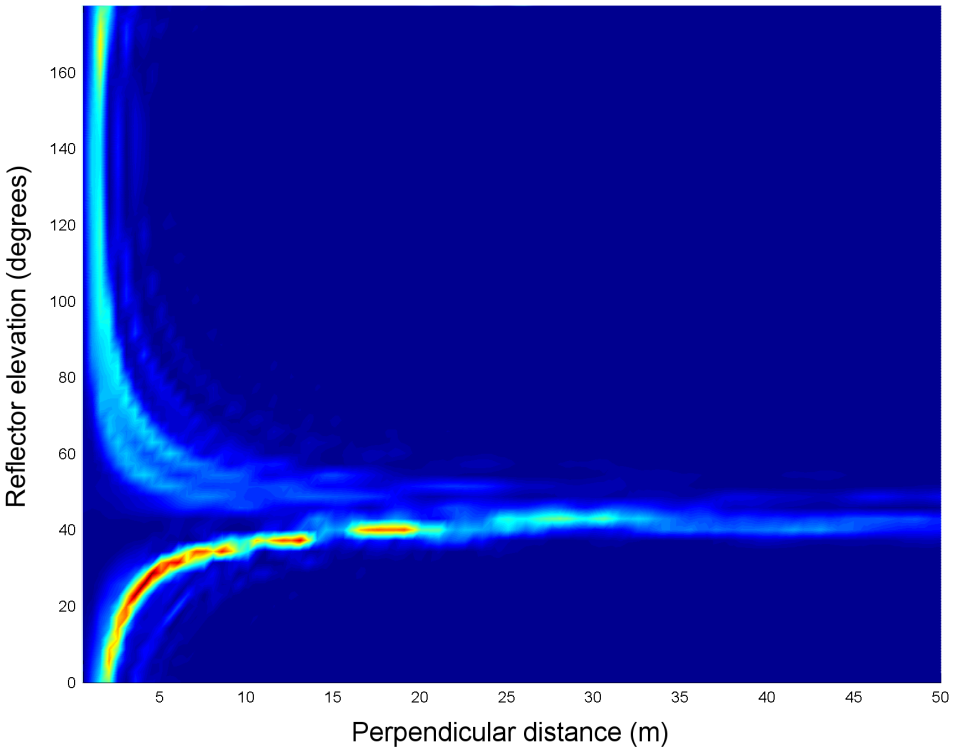


Figure 6.9: Shows the receiver output power in the frequency and time domain before and after interpolation

Canyon Environment PRN3

The simulation was executed for a canyon type environment. This is a very demanding environment for a GPS receiver as many satellites are out of view. The figures show that there is no direct line of sight for the satellite, however there are multiple reflected rays that reach the receiver. Figure 6.10 shows the outputs from the receiver in this environment for satellite PRN 3. The power output fluctuates rapidly from a full power signal to loss of lock. The ray-shooting algorithm outputs in Figure 6.11 at 5 and 10 minutes show up to four reflected rays from different positions reaching the receiver within the environment. In previous results the carrier interference between two rays has resulted in an oscillating power output. In this case there are four rays, each covering different distances causing a much more rapid and less predictable output. The power output stabilises at approximately a third of the power of a direct signal just before the signal is blocked by the canyon at 15 minutes. This is explained by Figure 6.11 (14 minutes), where only a single reflected ray reaches the receiver during this period, thus no interference occurs. The phase and positioning outputs from Figure 6.10 show significant

errors, where signal lock is rarely maintained. The pseudorange performs reasonably considering the conditions, averaging at about 30 m error but peaking to over 100 m for short periods.

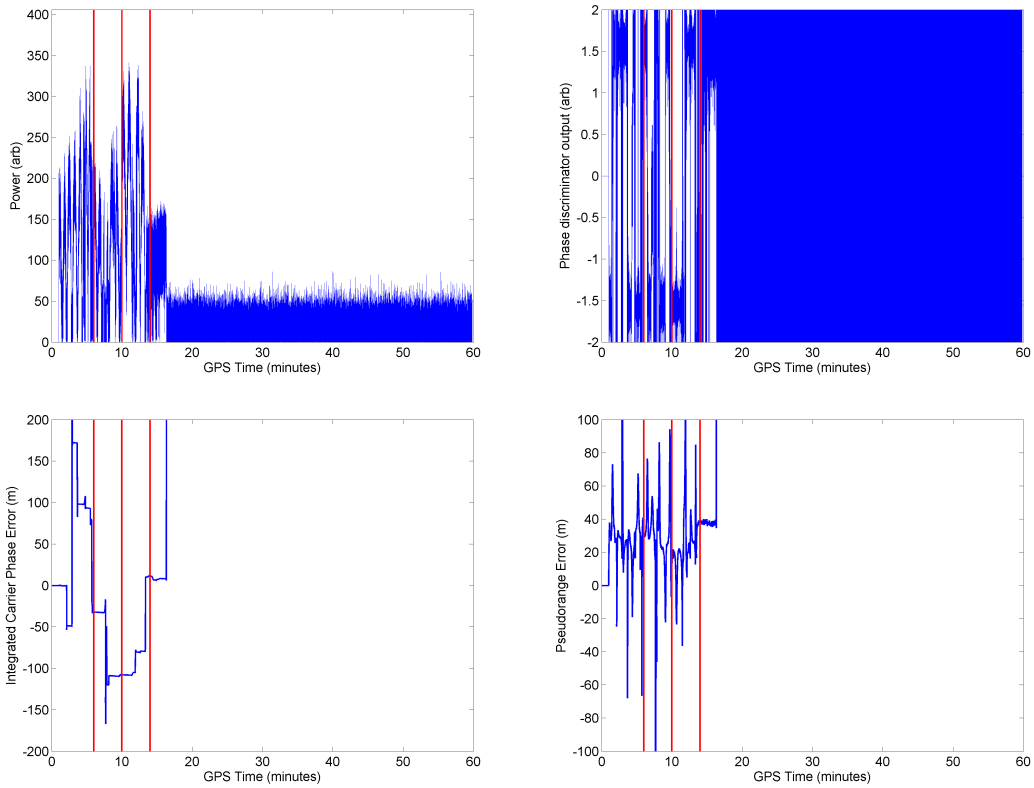


Figure 6.10: Receiver results from a canyon environment for satellite PRN 3.

Figure 6.12 shows the output from the MAT for this case. As mentioned above no direct rays reach the receiver so the receiver must track one of the reflected rays. This makes the MAT less viable as it can only relate the multi-path to the direct ray. However, a comparison of the distances covered by each ray relative to the ray being tracked is possible. By examining the output of the transform a peak is seen at 4 m and multiple peaks between 15 m and 20 m , no elevation angle is identifiable and it would not be helpful in this case. The first image from Figure 6.11 shows a blue and green ray travelling a similar path to the receiver. Assuming the blue ray is being tracked by the receiver, the green ray would cause the 4 m peak seen in the MAT. The red and green rays travel similar paths to each other but are significantly different from the blue ray causing two of the peaks between 15 m and 20 m . The second image from Figure 6.11 shows the red ray to have changed path slightly from the previous image, which may have caused the third peak between 15 m and 20 m .

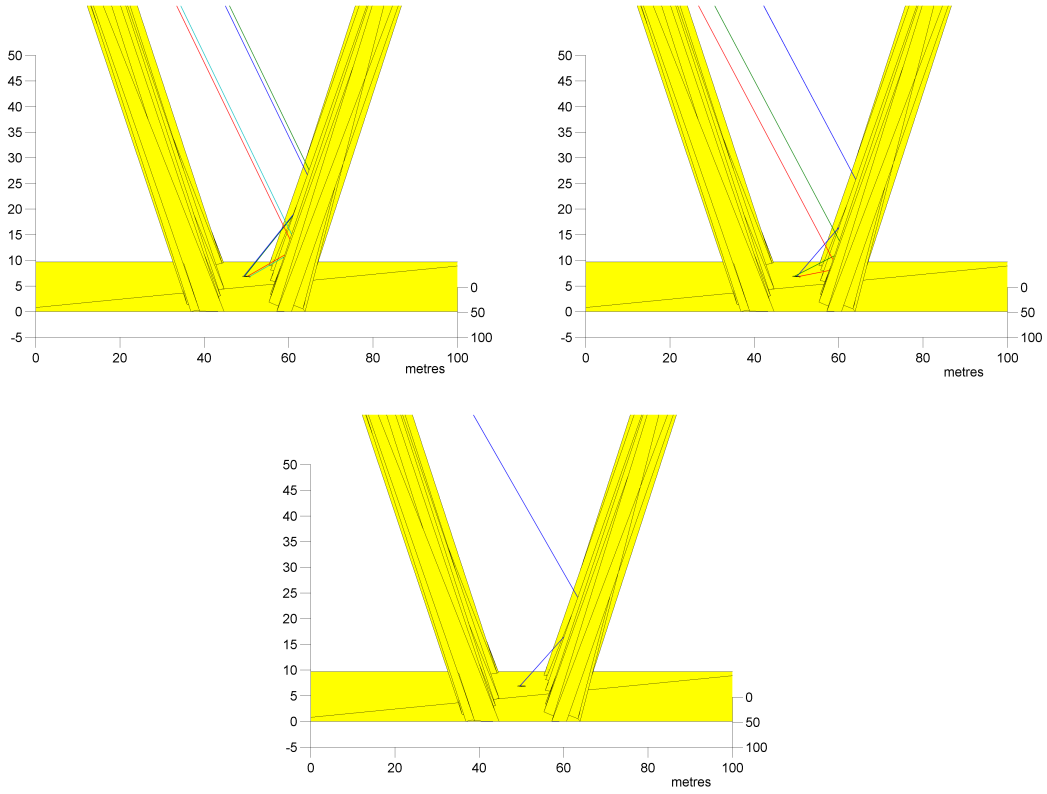


Figure 6.11: Environment and ray paths at 5, 10 and 14 minutes

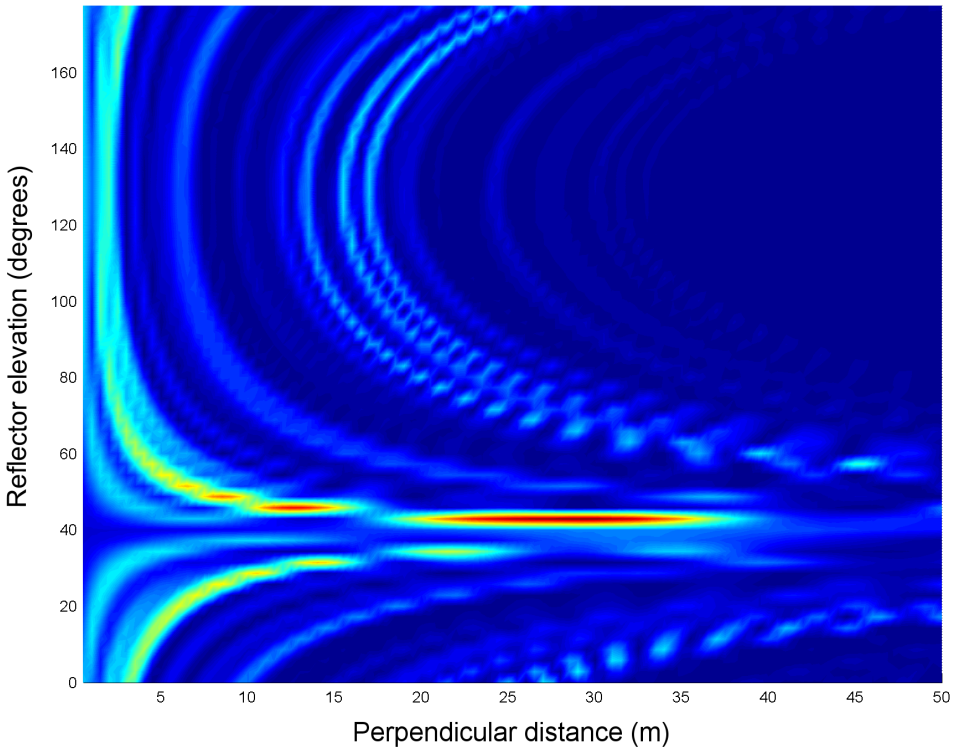


Figure 6.12: Shows the receiver output power in the frequency and time domain before and after interpolation

Canyon Environment PRN11

This simulation uses the same canyon environment as the previous one, however a different satellite signal is tracked. Figure 6.13 shows the power, carrier discriminator and positioning error output. The power outputs show fluctuations of a sinusoidal form similar to the results from the previous section. The carrier discriminator shows no loss of lock or cycle slips, except during the initialisation. The ICP shows a very small error, a half wavelength offset from a cycle slip near the start of the simulation and some very slight oscillations from the multi-path. The pseudorange shows a larger oscillating error from the multi-path as expected, varying between -5 m and +5 m . The power output shows two distinct oscillation periods, one has a small period (less than 1 minute) and is intermittent, the other has a longer period (10 minutes) and is stable but only occurs in the second half of the measurements.

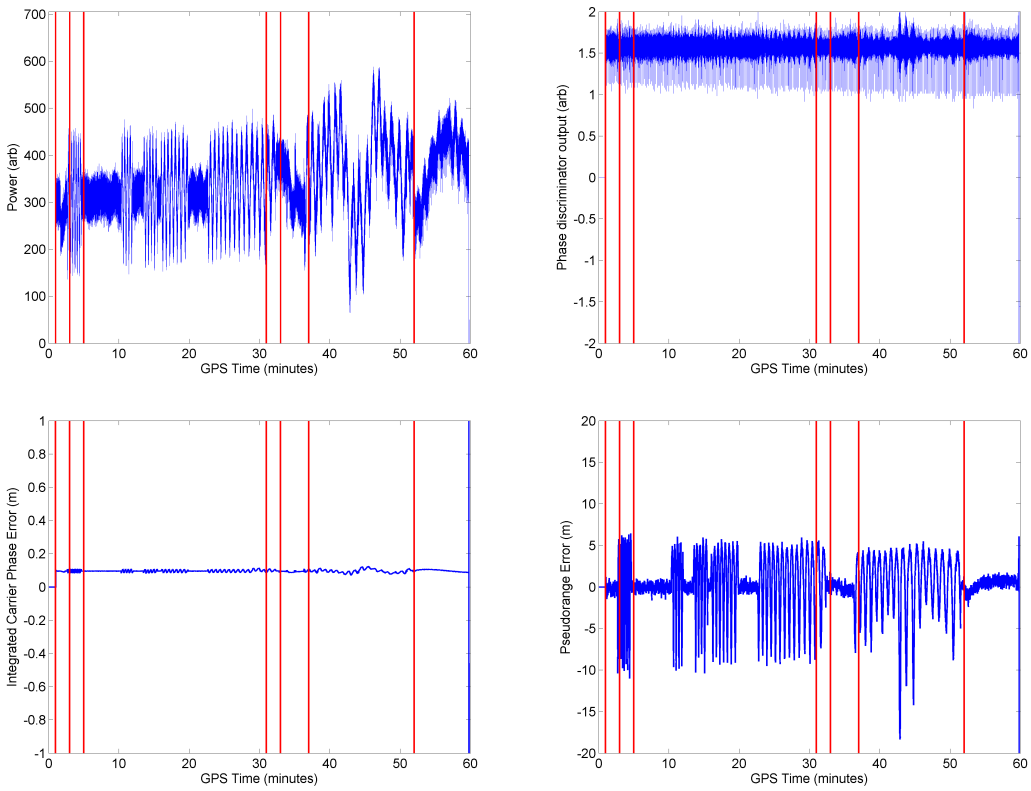


Figure 6.13: Receiver results from a canyon environment for satellite PRN 11.

To find the cause of the two interfering signals, the ray-shooter outputs from Figure 6.14 are used. The first three diagrams (1, 3 and 5 minutes) show before, during and after the first set of rapid oscillations. The diagrams show a single direct ray before and after, however during the interference, a direct and a reflected ray from the right hand canyon wall is shown. The next ray diagram is at 31 minutes where the power output shows the rapid oscillations and the start of the slower oscillations. The ray diagram shows two reflected rays showing that the right hand canyon wall causes the rapid oscillations, where as the left wall causes the slower oscillations. This corresponds to the previous work relating the period of the oscillating interference to the perpendicular distance of the reflector from the receiver. The diagram at 34 minutes shows only the left wall reflection, which is confirmed in the power output by the slow oscillation and lack of any rapid oscillation. The diagram at 38 minutes shows both reflections again causing both oscillation types in the power output and the final diagram only shows a left wall reflection corresponding to a slow oscillation. Despite multiple reflections and significant interference in this simulation the receiver had little problem maintaining lock on the signal and only small errors were shown in the ICP output. The pseudorange is not affected by the slower oscillations as the codes are not out of phase by enough, however the rapid oscillations cause errors of 5 m and -10 m .

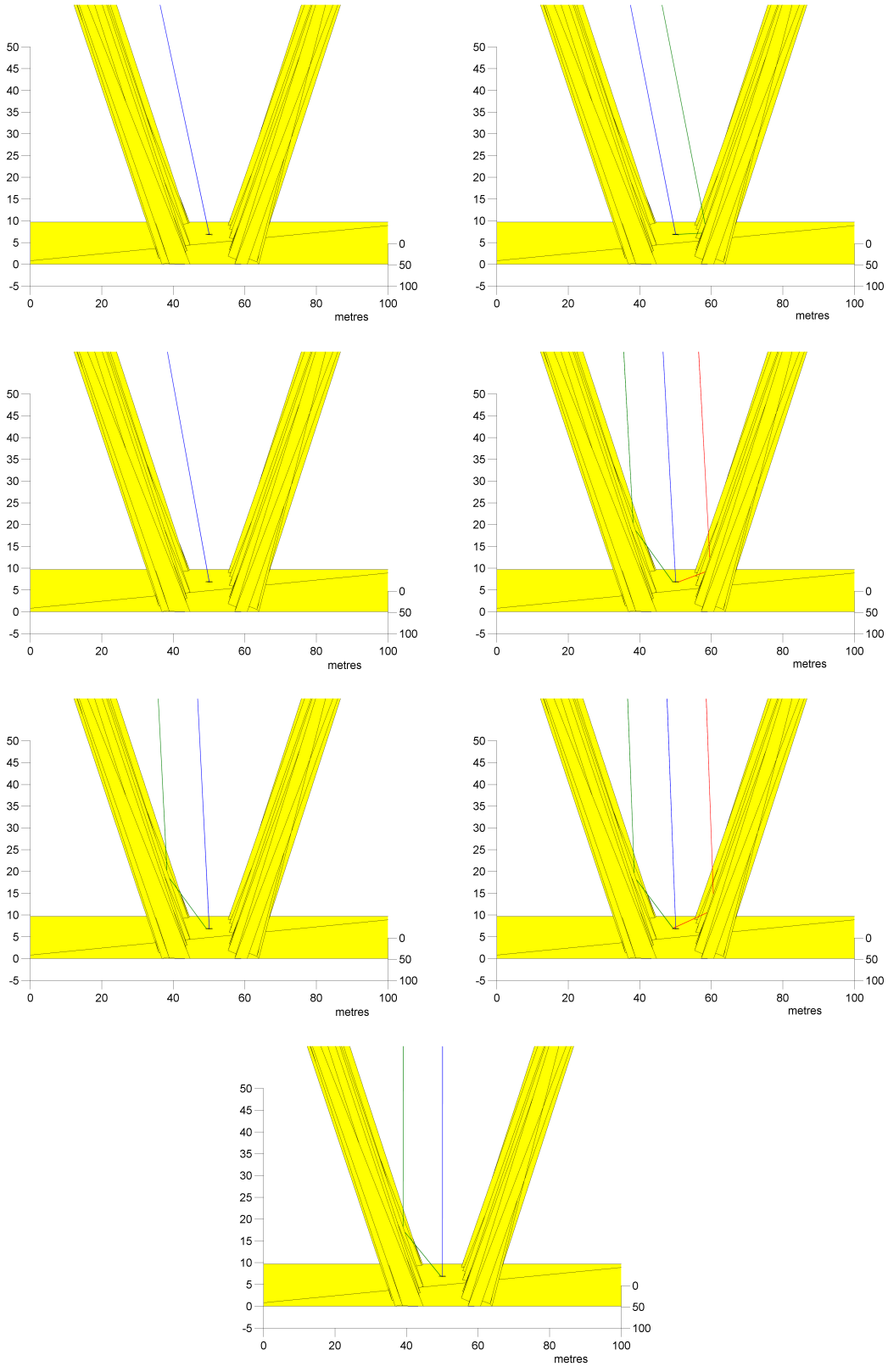


Figure 6.14: Environment and ray paths at 0, 3, 5, 31, 33, 36 and 52 minutes

The MAT in Figure 6.15 shows two multi-path signatures at 5 m and 35 m off reflectors at 20°. The distances can be verified to some extent by the two different interference periods shown in the power output in Figure 6.13. From the ray-shooter outputs in Figure 6.14 it can seen that two multi-path rays that reach the receiver, however what is not evident is that the satellite is at an elevation of 45° near parallel to the canyon. As mentioned previously this causes severe interpretation difficulties as the reflector elevation is not measured along the vertical plane but the plane connecting the receiver, the reflection and the satellite. This causes a steep reflector elevation of 60°, as shown in Figure 6.14, to be shown in the transform as a peak at 20°. However both multi-path signatures are clearly shown and with the satellite's position the path of the reflected rays can be determined.

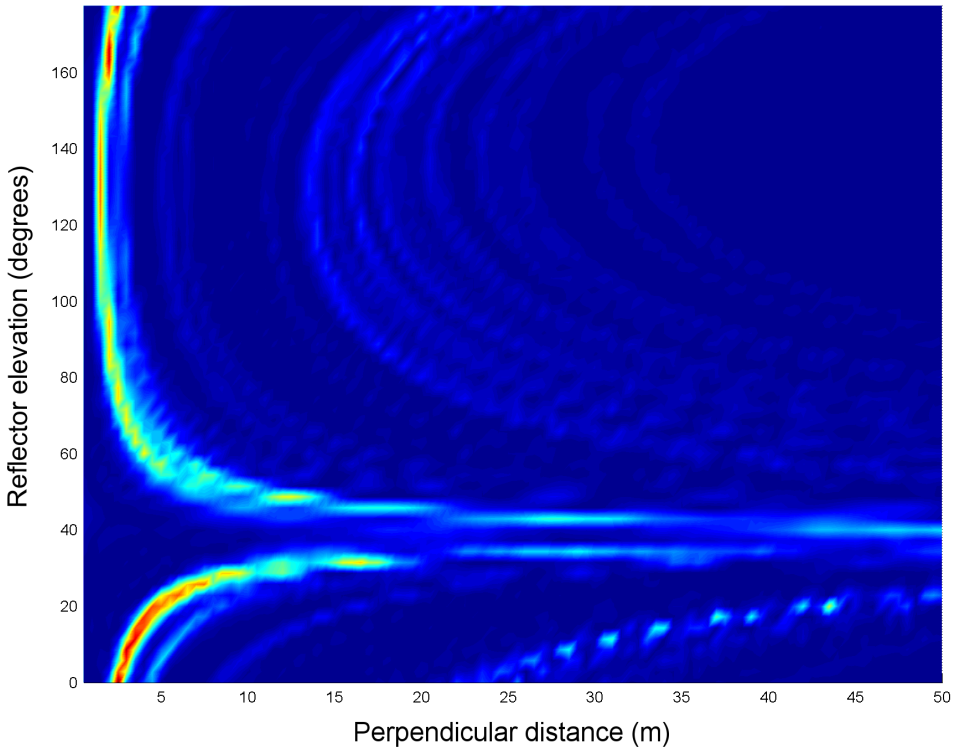


Figure 6.15: Shows the receiver output power in the frequency and time domain before and after interpolation

Hill Environment PRN11

This simulation shows similar characteristics to the previous one, where the distance between the reflection and the receiver determine the period of the carrier interference. Of particular interest are the very rapid oscillations shown in Figure 6.16 at 43 minutes.

These correspond to the long distance reflection shown in the third graph from Figure 6.17. This long range reflection affects the pseudorange more severely than the other nearby reflections as discussed in the previous section, causing an error of +/-20 m

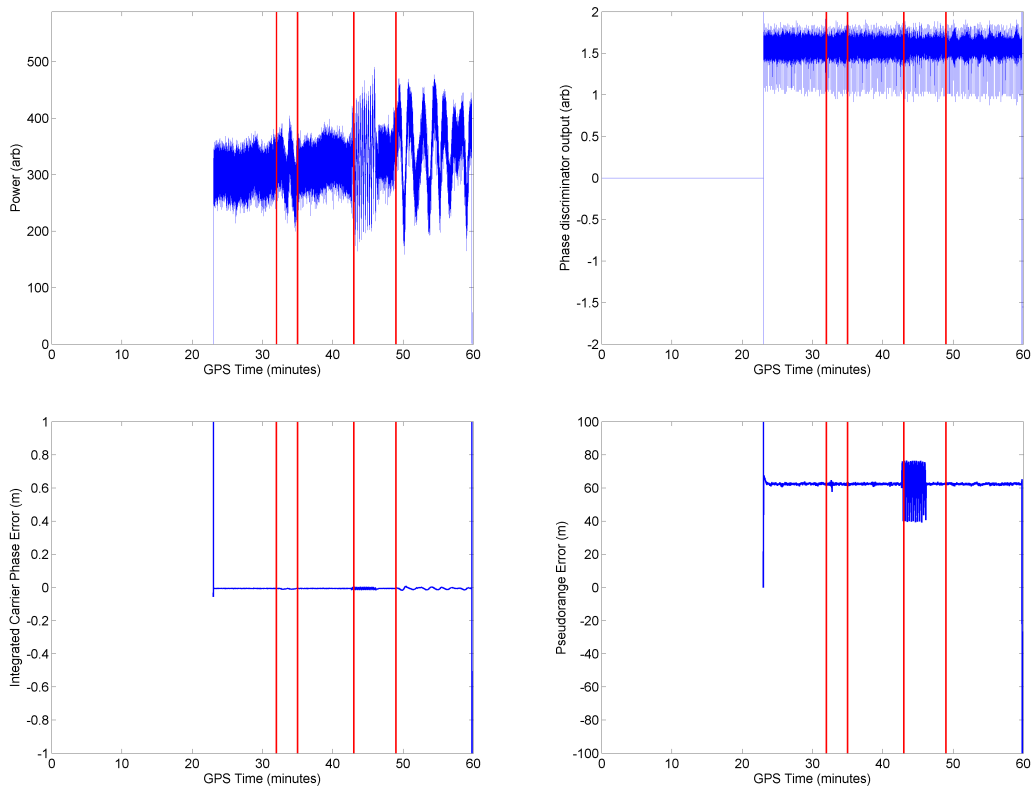


Figure 6.16: Receiver results from a canyon environment for satellite PRN 11

The MAT in Figure 6.18 shows two peaks at 8 m and 10 m as well as a faint peak at 50 m . The peaks shown are quite wide in elevation angle due to the short duration of each period of interference. However, three peaks can be matched to the three interference patterns shown in the power output in Figure 6.16. Where the 50 m peak is attributed to the very rapid oscillations in the power output at 45 minutes in Figure 6.16 and the 8 m and 10 m peaks are attributed to the slower oscillations at 50 and 33 minutes respectively. The ray-shooter outputs also confirm this with two nearby reflections and a distant reflection.

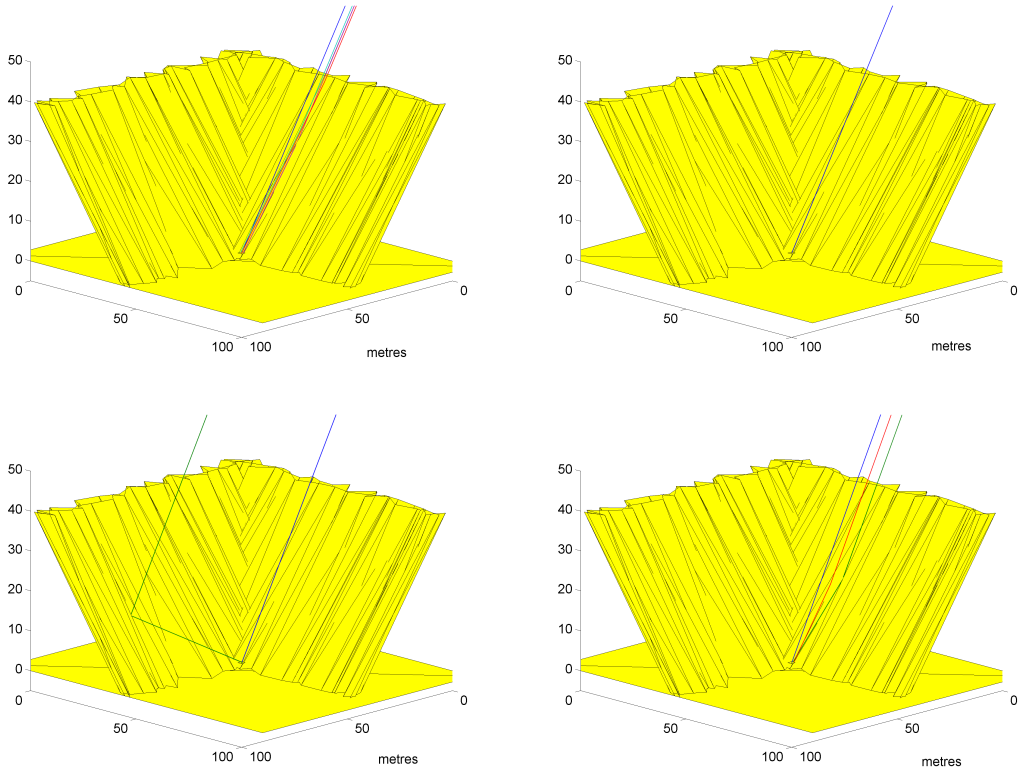


Figure 6.17: Environment and ray paths at 32, 35, 43 and 49 minutes

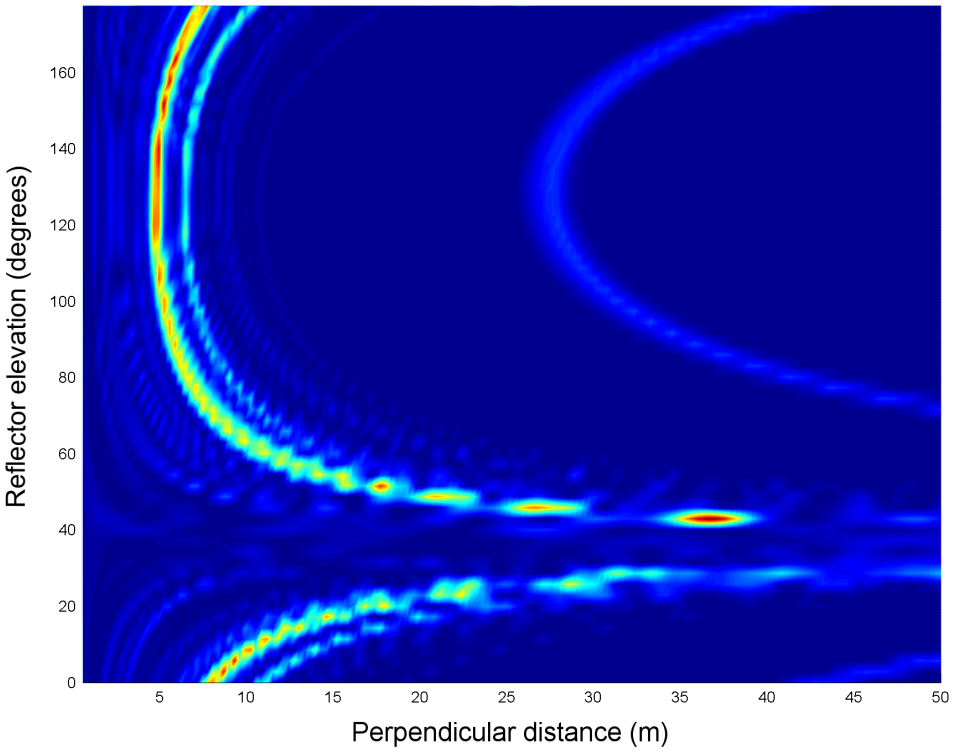


Figure 6.18: Shows the receiver output power in the frequency and time domain before and after interpolation

Urban Valley Environment PRN11

This simulation shows the ability of the receiver to track a very weak reflected signal without fully locking on to the signal. The output power in this case is barely distinguishable from noise and the carrier discriminator is constantly cycle slipping. The outputs from the ICP show an increasing error until the receiver is unable to maintain a lock and begins re-acquisition. The MAT is not shown for this scene as only one ray reaches the receiver and so no interference occurs.

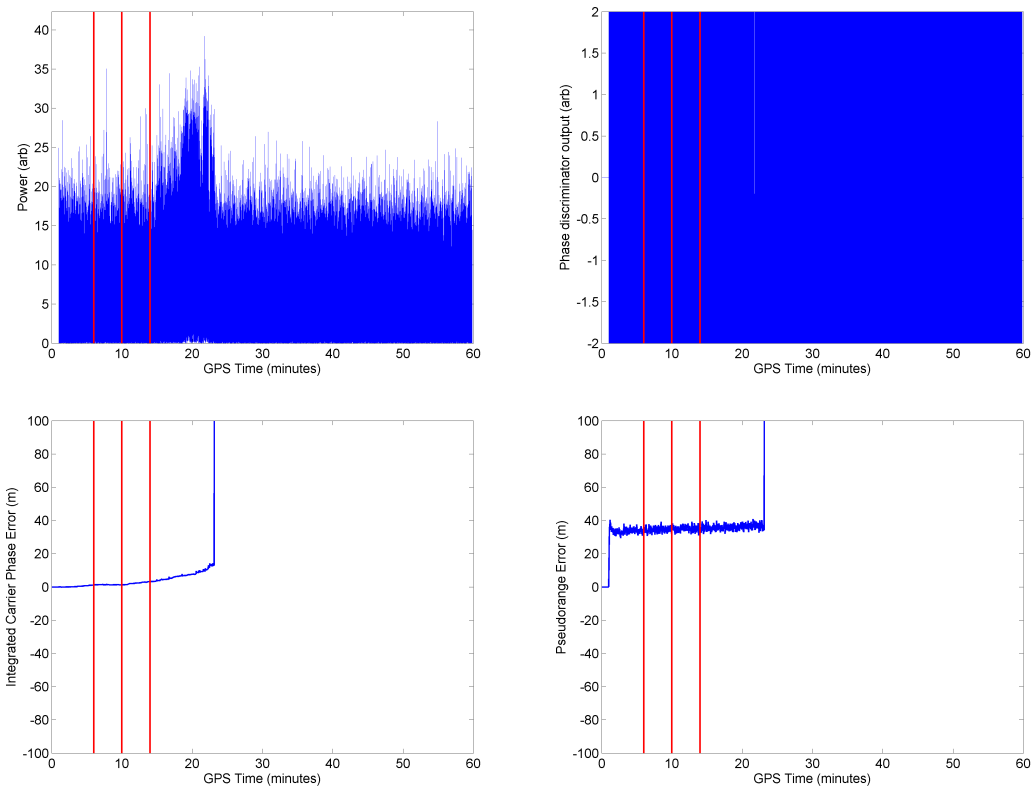


Figure 6.19: Receiver results from a dense urban environment for satellite PRN 11.

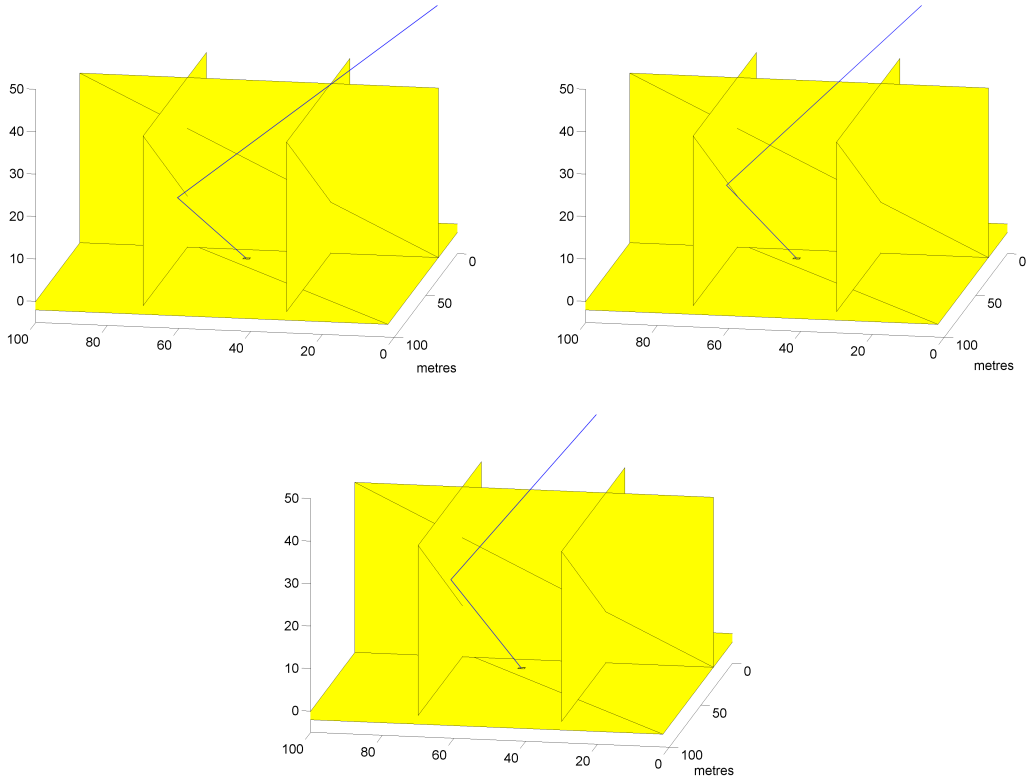


Figure 6.20: Environment and ray paths at 0, 15 and 30 minutes

6.3.3 Conclusion

The main conclusion from these simulations is that while a single multi-path ray is both easily identifiable and has little effect on the receiver tracking, multiple rays can severely disrupt the signal power and cause large tracking errors. A single reflected ray has been shown to cause an oscillating interference pattern, however multiple rays are shown to cause much more erratic patterns. This makes identifying the causes, even with ray-shooter outputs, very difficult.

A method to analyse multi-path effects using the receiver power output has been further tested in these simulations. The MAT has been able to separate the effects of different multi-path signatures and can determine how much of a deviation the reflected ray has undergone in comparison to the direct ray. As discussed through the results, interpretation of the outputs from the transform can be difficult as it assumes a two dimensional environment. When used for a 3-dimensional environment the transform determines the output for a two dimensional plane connecting the receiver, reflection and satellite positions.

An interesting factor found in these simulations is how the receiver copes under ex-

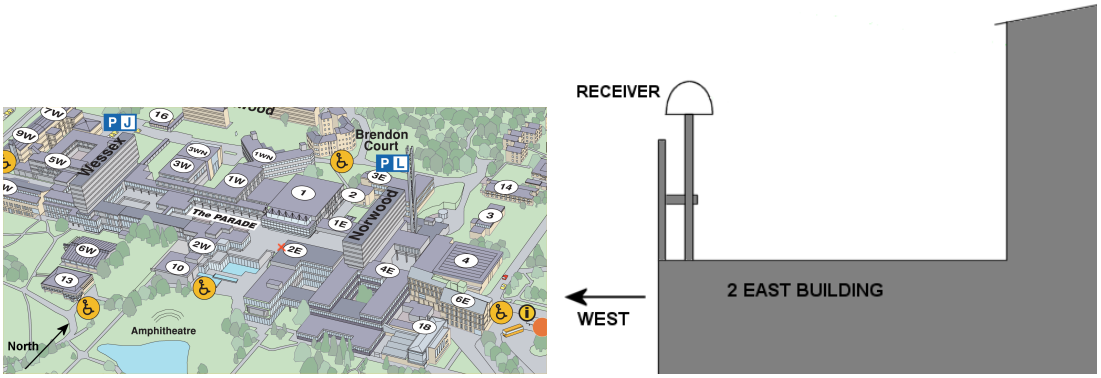
tremely difficult conditions. The experiment has shown the receiver to be very robust: only when the receiver was completely enclosed did a tracking failure occur. It should be considered that a real receiver also has ionospheric and clock errors to contend with and so may not be quite as robust as the simulated receiver shows.

6.4 Analysing Real Receiver Data

6.4.1 Introduction

This section uses the multi-path analysis transform (MAT) on real receiver data to determine its effectiveness. The power data from a receiver based at the University of Bath campus is analysed. There are multiple possible reflection points from buildings around the campus however the receiver is based at a significant height which reduces the possibilities of multi-path. Figure 6.21 (i) shows the environment around the receiver based at Bath University. The receiver is positioned on the roof of building 2E, to the west of a roof-top structure as shown in Figure (ii).

6.4.2 Results and Analysis

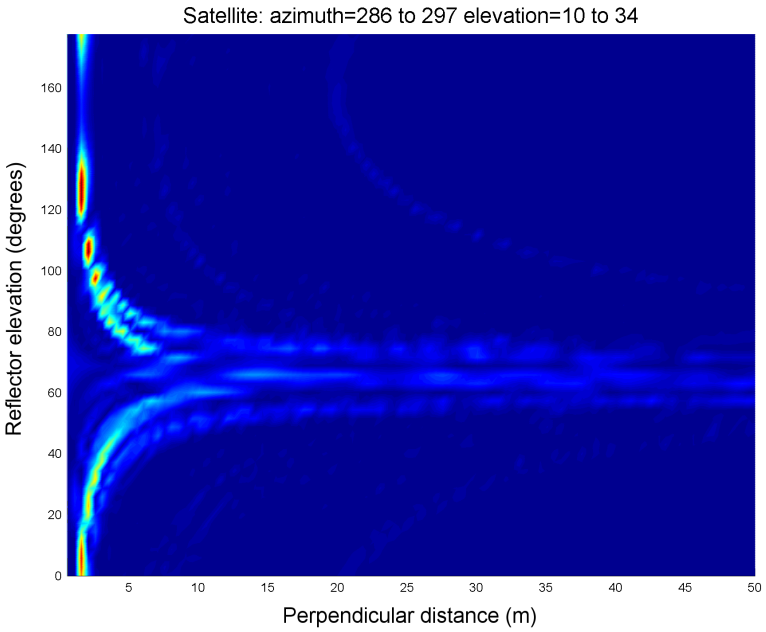


(i) (ii)

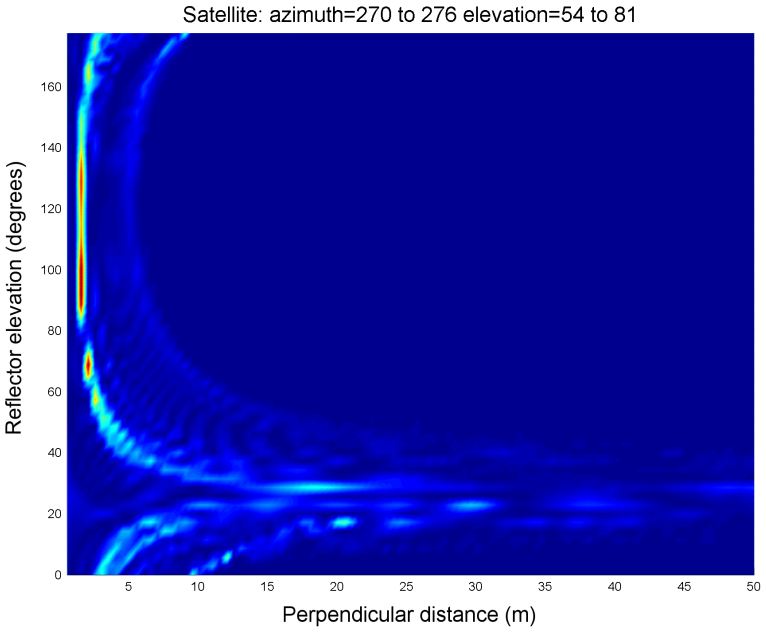
Figure 6.21: Figures showing the environment of the Bath based GPS receiver

Figures 6.22, 6.23 and 6.24 show the MAT for satellites to the west, south and east of the receiver respectively. It can be seen that the only multi-path signature occurs when the satellites are to the west of the receiver and that the multi-path is caused by a reflector approximately 2 m from the receiver. This signature matches with the roof-top structure shown in Figure 6.21 (ii). In Figure 6.22(i) two peaks are shown at 0 ° and 120 °,

which may be due to the ray reflecting off the roof and the roof-top structure, see Figures 6.25(i). This cannot be accurately analysed by the transform as it assumes a single reflection, however in this case where the walls are 90° apart, it can be seen from Figure 6.25 that the second reflection has no effect on the distance measurement. The rays shown in Figure 6.25 are those determined by objective analysis of the transform output but cannot be proved correct. However, it is highly unlikely that the oscillating interference that varies frequency with satellite elevation could be due to the ionosphere or any other error sources. Figures 6.23 and 6.24 show little evidence of multi-path as the roof top structure would reflect rays away from the receiver from these angles. Some weaker multi-path signatures were found from satellites between azimuth angles of 240° to 300°, but the strongest are shown in Figure 6.22.



(i)



(ii)

Figure 6.22: Multi-path transform for PRN 10 and 11, west of receiver

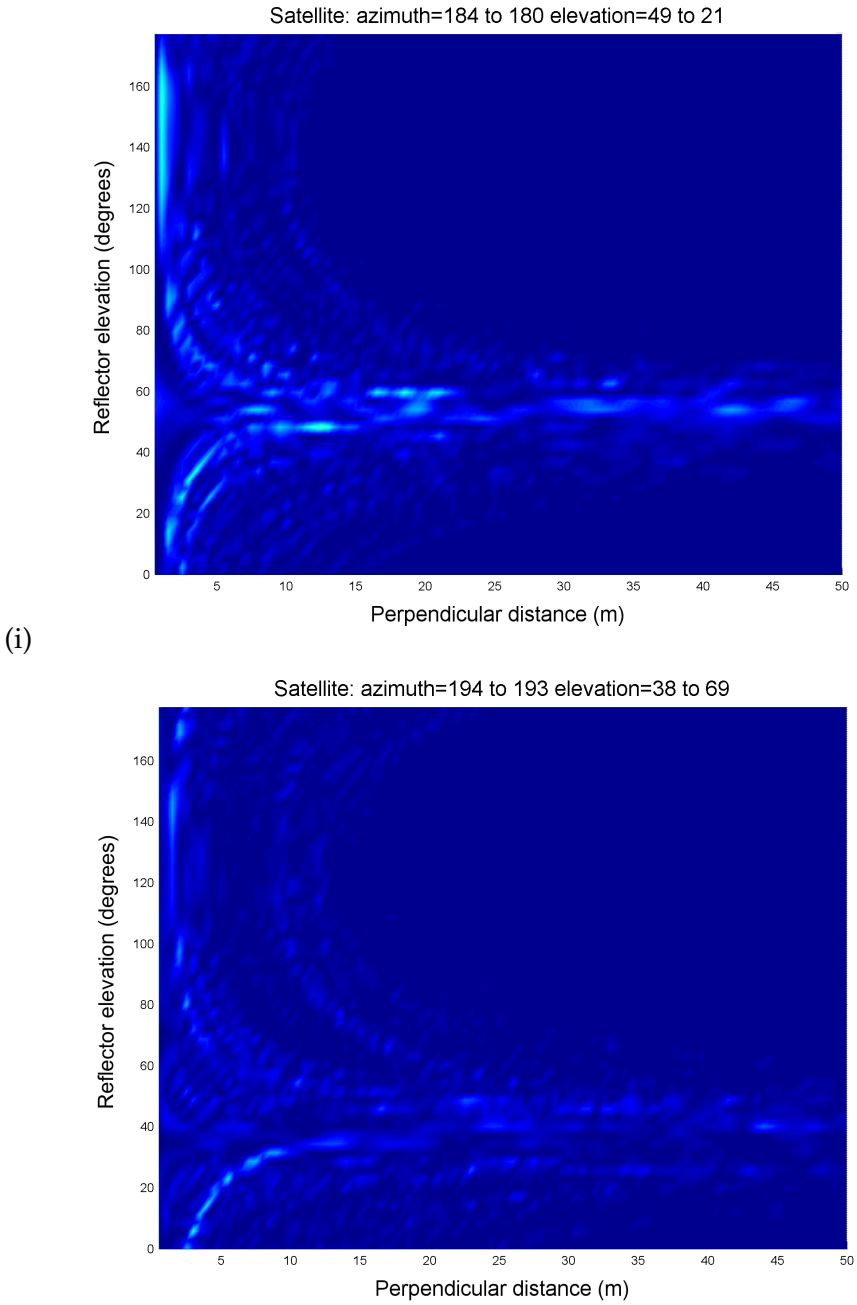


Figure 6.23: Multi-path transform for PRN 10 and 13, south of receiver

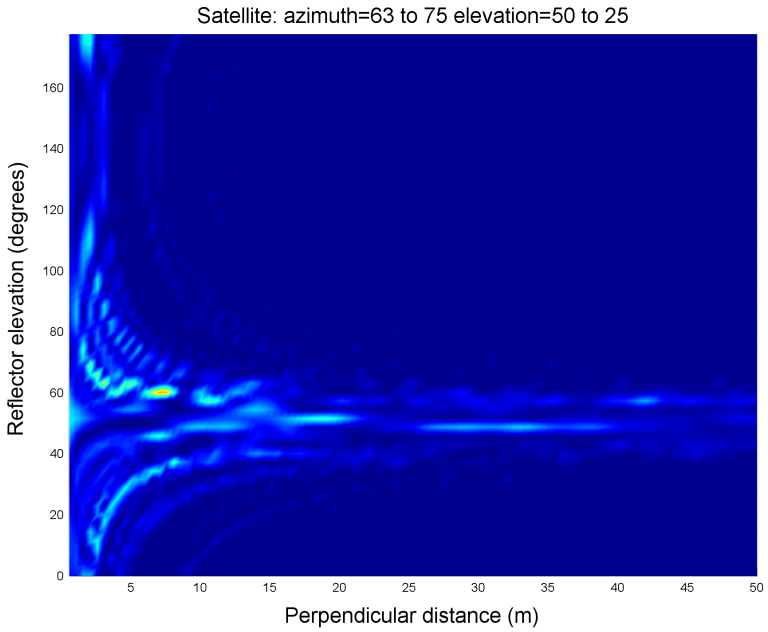
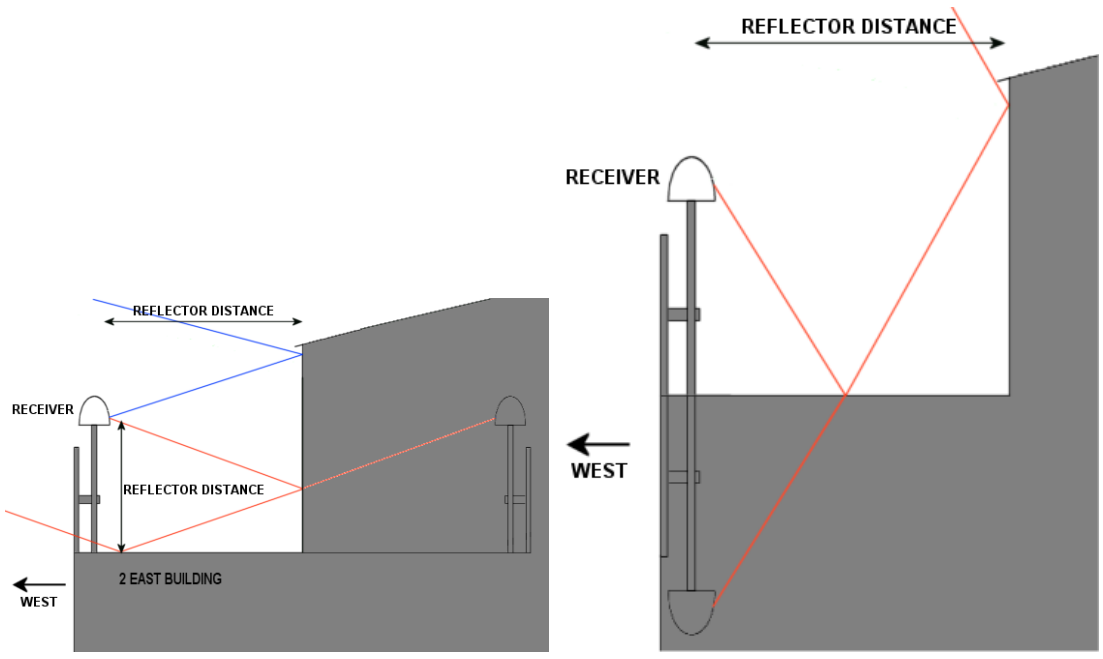


Figure 6.24: Multi-path transform for PRN 13, east of receiver



(i) (ii)

Figure 6.25: Figures showing the environment of the Bath based GPS receiver

6.4.3 Conclusion

This experiment has shown that the multi-path analysis transform works on real receiver data. Further testing is required under a more controlled environment, where the number of multi-path sources can be limited.

6.5 Summary

This chapter has investigated the oscillating carrier interference observed in the receiver power outputs when under multi-path conditions. A method to identify and analyse the multi-path properties has been derived and tested on simulated and real receiver data. The multi-path analysis transform has performed well in all but the most difficult conditions. The simulator has shown that it can model large simple environments, but may have issues when modelling large complex environments. This problem is believed to be solvable with some further development.

Chapter 7

Case Study of Tromsø based receiver

7.1 Introduction

In this chapter an isolated GPS signal fade event is analysed to the determine the cause. It was observed on three closely located GPS scintillation receivers in northern Norway, Figure 7.1. The event lasted about 10 seconds and was identified because the GPS receivers were configured to collect fast phase and amplitude data at 50 Hz , rather than the usual 1-minute index of scintillation, S_4. In fact the S_4 index was not particularly high at 0.4, but the signal amplitude showed a short rapid fluctuation that was repeated on all three GPS receivers. The event was studied using the multi-path tools developed and other instrumentation showing the local ionospheric conditions.

7.2 Multi-path Analysis

7.2.1 Introduction

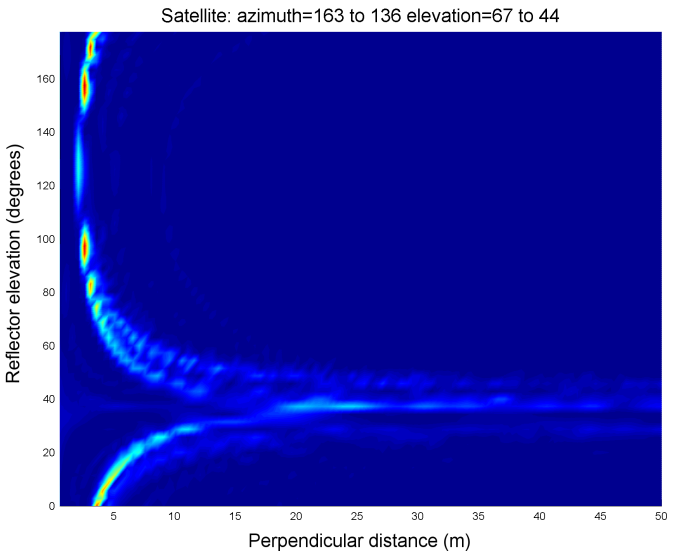
The methods and knowledge obtained in the previous multi-path chapters is used here to determine whether the fade in the GPS receiver power could be caused by multi-path. The signal is analysed using the MAT derived in the previous chapter and simulated to determine the signature of any possible multi-path.

7.2.2 Analysis of real data using the multi-path analysis transform

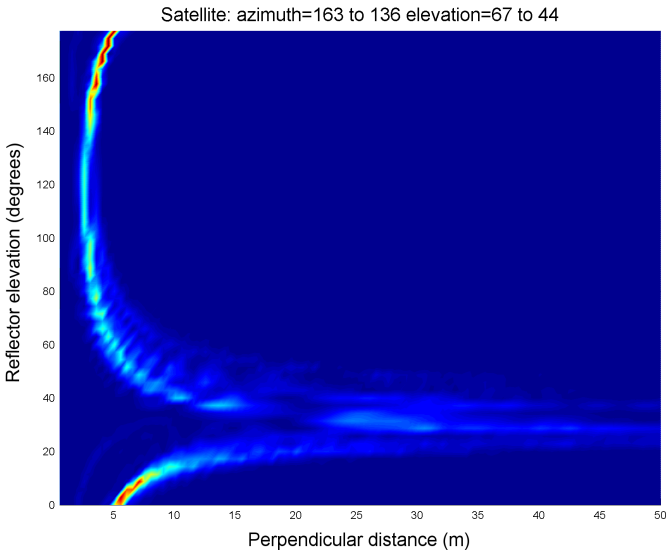
The MAT was ran on the power outputs from the three receivers from 12:00UT on the 7th until 02:00UT on the 8th November 2004 to determine if any multi-path signatures were present. Figure 7.2 shows the peak outputs found during the period analysed for

the two locations. Figure (i) shows two peaks at 100 ° and 160 ° . The 100 ° peak cannot be associated with a reflected ray as there are no vertical reflectors 5 m from the receiver. It may be due to a double reflection causing an erroneous or difficult to interpret elevation value. The 160 ° peak matches with the 20 ° slope of roof of the building the receiver is mounted to. Figure (ii) shows a clear signature of a ground reflected multi-path ray with a distance of 5 m which matches the mounted height of the receiver. The Figures in 7.3 show the outputs during the time of the event in question. The only multi-path signature is weak and matches the ground multi-path seen in Figure 7.2(ii), which is of a much lower frequency than the rapid fade. The transform was extended to 250 m to analyse more rapid multi-path fluctuations, however no signatures were seen.

Figure 7.1: Pictures of the GPS receivers based at Tromsø



(i)



(ii)

Figure 7.2: MAT output for satellite PRN 5

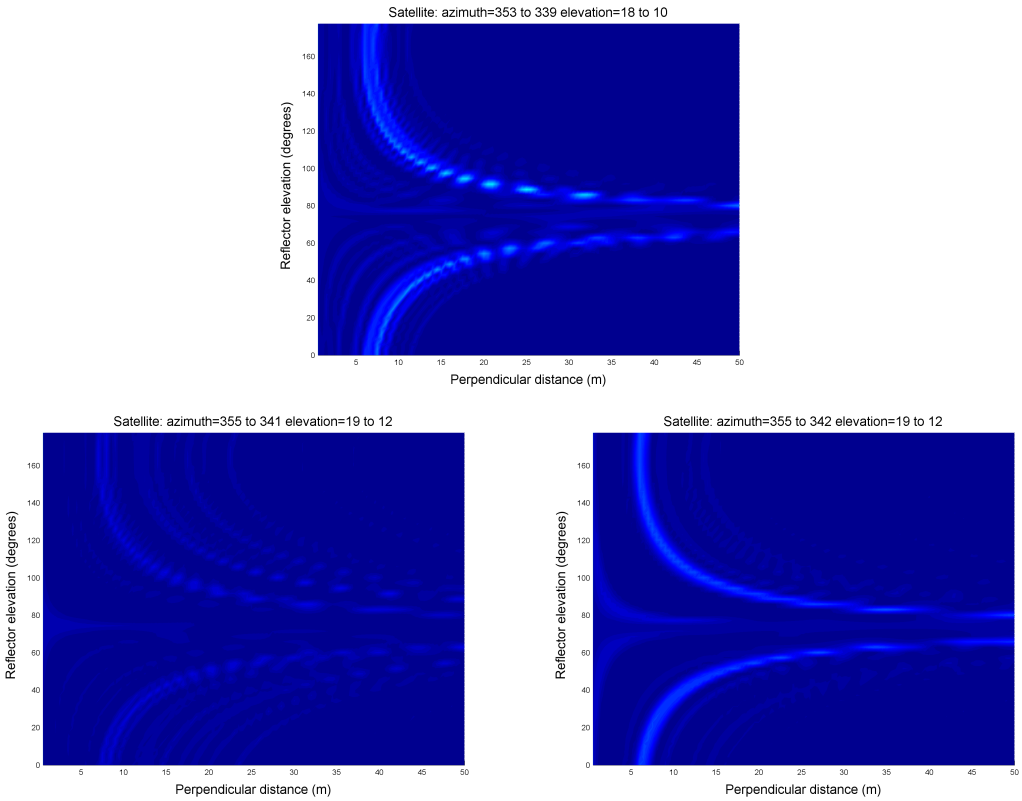


Figure 7.3: MAT output for the period during the fade

7.2.3 Simulation of the local environment

The simulation is now used to confirm that the signatures found can be attributed to the local environment. It also shows that nothing nearby the receiver could cause the rapid fade event. A number of simulations were processed to investigate the environment around the receiver. Three possible sources of multi-path were identified and simulated to determine whether they could cause the rapid fade seen in the real receiver data. The first was the EISCAT radar, which has a large cylindrical antenna measuring 40 m in height and 120 m in length. This structure is approximately 200 m from the receiver and therefore is unlikely to cause multi-path unless the satellite is at very low elevation. The second source considered is the ground, where a low elevation reflection could reach the receiver avoiding the radome base. The last candidate is the building that two of the receivers are mounted on, however if this was the cause the fade would only be seen on two of the receivers. Figure 7.4 shows the environments modelled by the simulator. Positions and sizes of the structures were obtained from a plan of the EISCAT facility courtesy of the University of Tromsø.

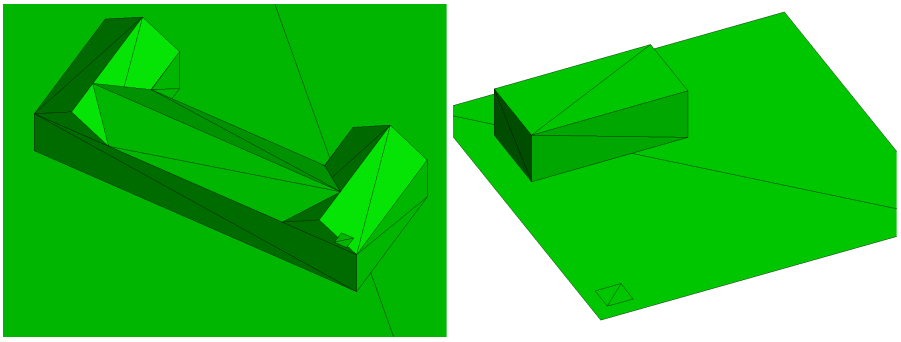


Figure 7.4: Modelled environment of the Tromsø based receivers

7.2.4 Simulation results

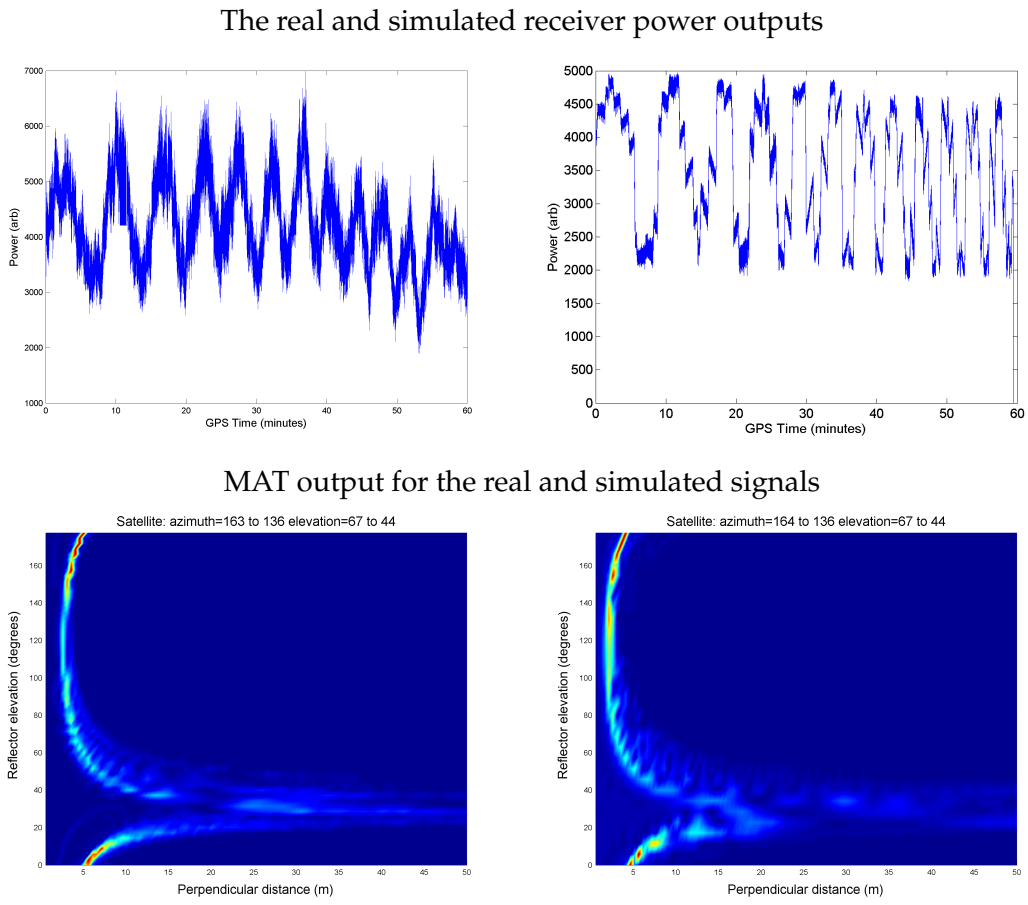


Figure 7.5: The signals and the MAT output for the ground reflected multipath case

The results from the EISCAT radar simulations showed no interference on the signal from satellite PRN 5. The radar is too distant from the receiver and in a similar direction as the satellite. Despite a number of simulations varying size and positions of the

structure no multi-path was observed.

Figure 7.5 shows the power output from the real and the simulated data, where the receiver is mounted 5 m above the ground. A very strong similarity can be seen in the power fluctuations where the period of the oscillations reduces as the time progresses. The number of oscillations in the one hour period of measurements from the real and simulated cases also match and show similar changes in frequency. Figure 7.5 also shows the real and simulated outputs analysed using the multipath analysis transform. It clearly shows the same peak in the real and simulated results representing a reflection of a horizontal surface at a perpendicular distance of 5 m .

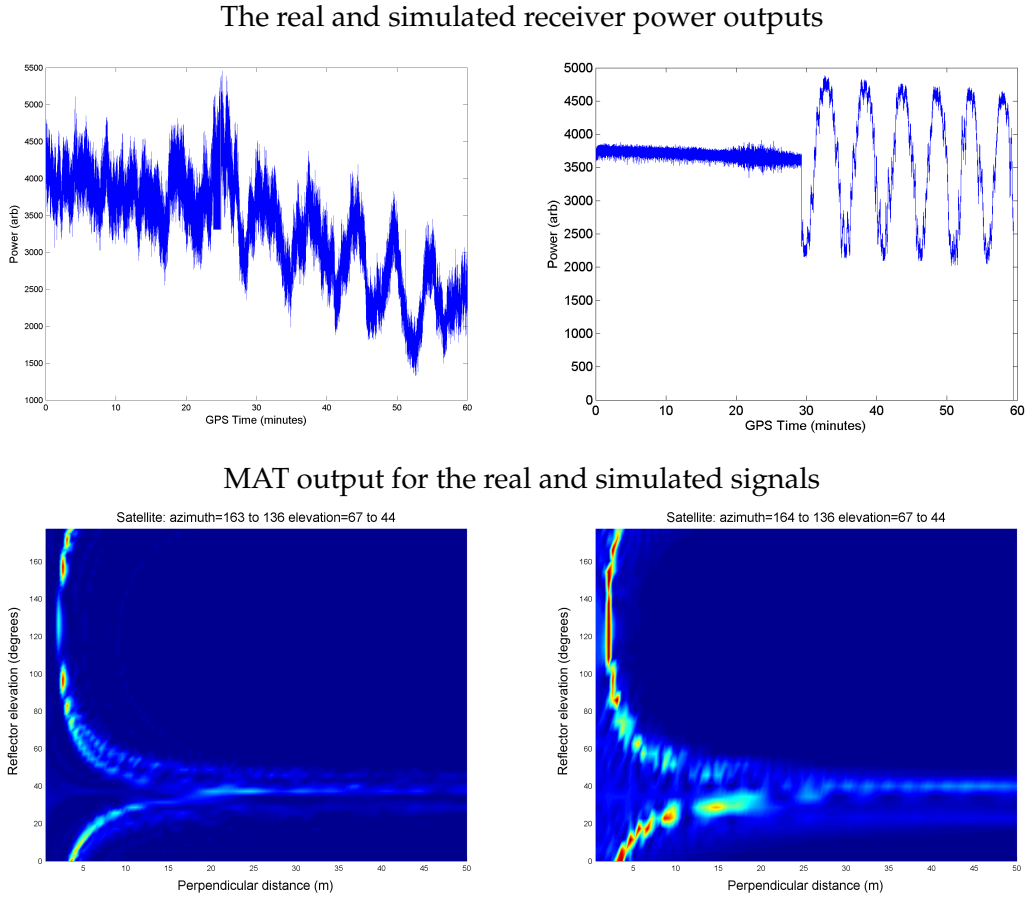


Figure 7.6: The signals and multi-path analysis of the roof reflected multipath case

Figure 7.6 shows the power output from the real and the simulated data, where the receiver is mounted 1m above the building. A similarity can be seen in the power fluctuations where the period of the oscillations reduces as the time progresses. Figure 7.6 also shows the real and simulated outputs analysed using the multipath transform. It clearly shows a similar curve in the real and simulated results. However, the simulated results show no specific elevation peaks which may be due to small inaccuracies in the building model.

7.3 Ionospheric experimental equipment and analysis

Northern Scandinavia is located under the auroral oval and hence is a region in which auroral scintillation is likely to be encountered. There are a number of instruments for ionospheric research in that region, clustering around the European Incoherent Scatter (EISCAT) radar facility. Of particular interest to this study are instruments that indicate the presence of auroral activity: magnetometers and all-sky cameras (ASC).

Three GPS scintillation receivers have been deployed at the European Incoherent Scatter (EISCAT) radar transmitter site (69.6 ° N, 19.2 ° E) near Tromsø, Norway (Figure 7.7) since June 2004. They form part of a network of GPS receivers in northern Europe specially deployed to monitor the effects of ionospheric irregularities on GPS signals. One of the Tromsø based receivers is a Novatel OEM 4 receiver, specifically configured to record the phase and amplitude of the L1 signal simultaneously on multiple GPS satellites, providing samples at a rate of 50 Hz (*Van Dierendonck et al.*, 1993). Co-located with this receiver is a second receiver developed by Cornell University measuring only the amplitude of the L1 signal at 50Hz (*Beach and Kintner*, 2001). The antennas for these two receivers are about 1m apart (east-west). A third scintillation monitor (of the Cornell type) is located at the Heater transmitter site (denoted HTx) around 240 m to the west of the others.

The IMAGE chain of magnetometers (*Viljanen and Häkkinen*, 1997) operates in Fennoscandia covering roughly 2 hours in MLT and magnetic latitudes 55 ° -75 ° . The meridional middle chain of IMAGE is particularly suitable for electrojet studies. In longitudinal direction the coverage is best in the standard auroral oval region (64 ° -67 ° magnetic latitudes). Currently, the network includes 29 magnetometers with sampling rate of 10 s .

The Finnish Meteorological Institute (FMI) operate eight all-sky cameras (ASC) as part of their MIRACLE network (Magnetometers, Ionospheric Radars All sky Cameras Large Experiment). The analysis presented here uses data from Kilpisjärvi (69.02 ° N 20.79 ° E), see Figure 7.7. The data are sufficiently accurate to observe global auroral movement and auroral arc formation. The current cameras operate with a filter wheel capable of accommodating five filters. Three filters are currently used with centre wavelengths of 427.8 nm (blue), 557.7 nm (green) and 630.0 nm (red). The ASCs make continuous observations updated every 20 s at 557.7 nm and every minute at 427.8 nm and 630 nm . In this study where we are predominantly interested in the dynamic auroral behaviour, we make use of the more frequently sampled 557.7 nm images. Once the images have been the collected, a brightness calibration can be added to derive absolute radiance.

Figure 7.7: Locations of the all sky cameras and the GPS receivers.

The radiance images can be used in two ways. Firstly, if a particular auroral feature can be observed from two or more ASCs at different locations, the exact altitude can be determined by triangulation. The images then can be projected into geophysical (or geomagnetic) coordinates to determine the auroral structure position. Secondly, if no significant commonalities exist, as could well be the case with significant separation between sites, the default assumption of 110 km altitude is used.

The Kilpisjrvi IRIS imaging riometer in northern Finland (69.05 N, 20.79 E) shows radio signal fading at 38 MHz and around the time of interest showed evidence of strong absorption. Further details about the imaging riometer can be found in Senior and Honary, (2003). This fading at low frequencies is likely to be associated with absorption at D-region altitudes which implies energetic particle precipitation, causing low-altitude ionisation enhancements. However, it is noted that scintillation on the GPS signals shows evidence of both increases and decreases in signal strength, which is indicative of diffraction not absorption and is not necessarily associated with ionisation at the same altitudes affecting the riometer.

7.4 Ionosphere Analysis

7.4.1 Introduction

The ionosphere can affect the integrity of a GPS signal. The effects are mainly geographically localised to the equatorial and the auroral regions. They are solar-cycle dependent, particularly at the equatorial regions, where the greatest effects occur at solar maximum (*Basu et al.*, 1988). Scintillations can occur at mid-latitudes, for example *Ledvina et al.* (2002) report amplitude scintillations of magnitudes greater than 20 dB,

sufficient to degrade or disrupt GPS signals, at magnetic latitudes of 53 ° N over the north-eastern USA. In the northern high-latitudes, scintillations have been observed on GPS signals at Fairbanks, Alaska (*Doherty et al.*, 2000) and from Svalbard (*Basu et al.*, 1998). GPS L1 frequency (1575 MHz) observations of scintillation in the polar cap are probably associated with irregular structure within polar-cap patches (*Essex et al.*, 2001). Such convecting patches seem to be the source of scintillation during the storm of October 2003 (*Mitchell et al.*, 2005). Further information about GPS scintillations can be found in *Basu et al.* (2001).

7.4.2 Results and Analysis

During the period of interest on 8 November 2004, a significant amount of fading occurred at each GPS receiver on the link to satellite PRN 5 towards the north-west. The elevation of the satellite was about 15 ° . The fade had a very short duration of approximately 10 s and occurred at around 01:23:20 UT. The data showed very similar fade patterns at each receiver, with a time offset in the occurrence between the co-located receivers and the other receiver based to the west. It was possible that the cause of the scintillation was a moving phenomenon, not two separate occurrences. As there are effectively only two receiver locations (two are in close proximity) only a single velocity component could be calculated. This was calculated by cross-correlating the fade pattern from each set of receiver data to determine the time offset between sites. The position and velocity components of the signal fades in the receiver data were projected along the satellite-receiver path onto the ionosphere to find the east-west velocity component at a range of hypothetical shell heights (for shells of constant altitude above the Earth).

Figure 7.8 shows the fading pattern in each set of GPS receiver data. The upper and middle plots show the raw power measurements for the co-located receivers and it can be seen that almost identical patterns of fading occur. The signal amplitudes both increase above and decrease below the nominal signal amplitude which is characteristic of diffractive scintillations. The lower plot shows a time series of signal strength from the same one-minute period for the second, separated, Cornell receiver, revealing that the onset of the event was approximately 0.3 seconds earlier. Thus receivers located some 240 m apart at HTx and HRx exhibit similar fading patterns with a time offset.

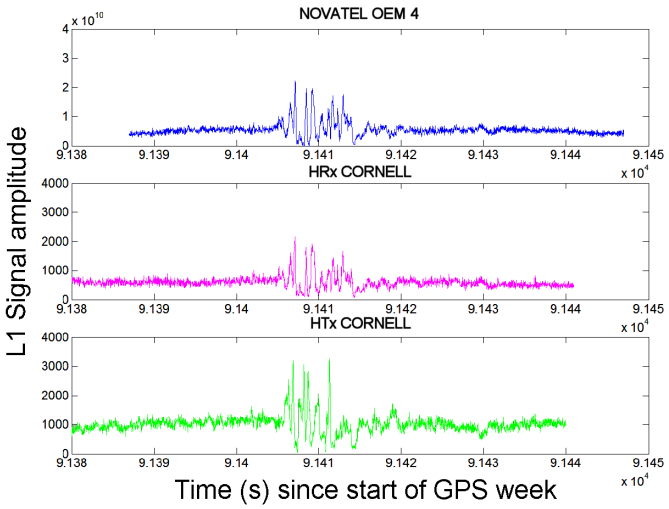


Figure 7.8: Signal amplitude measurements from the GPS receivers. The time interval shown is from 01:23:00 to 01:24:10 UT on 8 November 2004

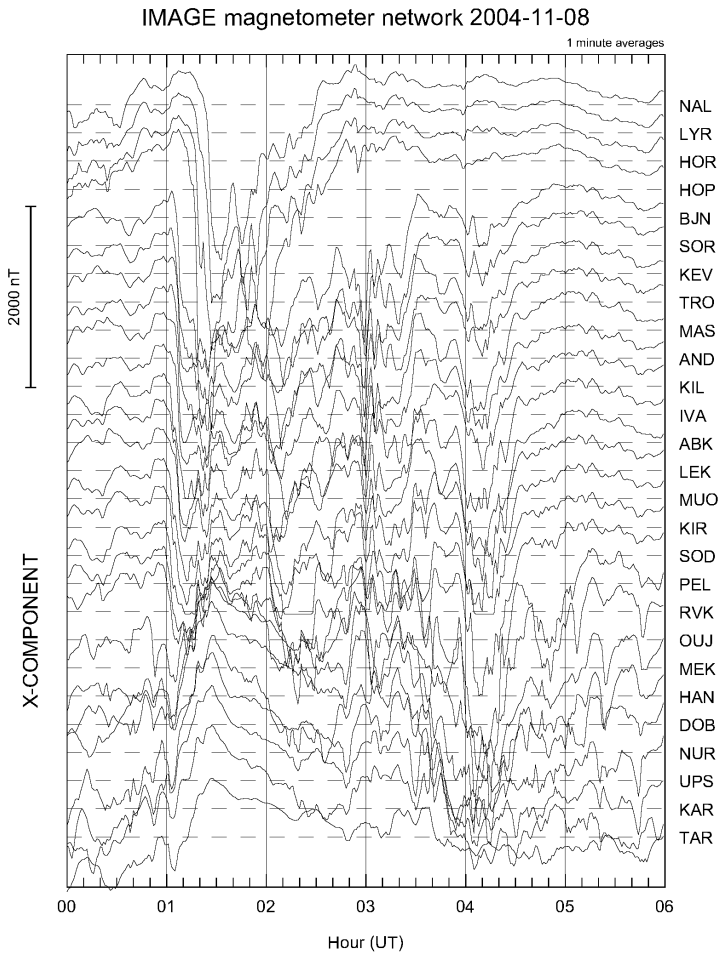


Figure 7.9: The x component of the magnetometer deflections during 00:00 to 06:00 UT on 8 November 2004

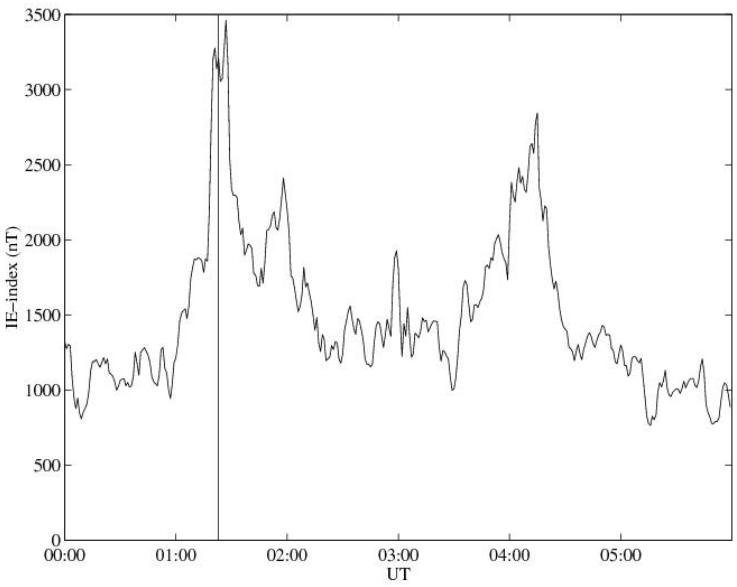


Figure 7.10: The IE index for 01 to 06 UT on 8 November 2004.

The geomagnetic conditions were disturbed at this time with the interplanetary magnetic field southward for several hours. Figure 7.9 shows the x component of the magnetometer deflections from IMAGE during the period of interest. Strong deflections can be observed around the time of interest; 0123 UT. Figure 7.10 shows the IE index, a measure of electrojet activity (*Kauristie et al.*, 1996), derived from the IMAGE magnetometer data for 00 to 06 UT, with the event at 01:23 UT marked.

The ASC images were examined for auroral structures. Some of the ASC data were degraded by cloud cover for these times, however all showed significant auroral activity, particularly visible at 557.7 nm . The ASC images from Kilpisjarvi at approximately 01:23 UTC are shown in Figure 7.11. In the image shown here the ASC were projected onto a specified shell of constant altitude of 110 km , assuming that the emissions come from atomic oxygen. A moving auroral feature was identified around the same time as the GPS event.

Starting from the assumption that the GPS and the all-sky camera events resulted from a moving auroral arc at the same altitude, the altitude and velocity were calculated as follows. Assuming that the precipitation event occurred at an altitude ranging between 80 km and 300 km the corresponding velocity range calculated from the time offset between the two ground-receiver events is between 758 ms^{-1} (80 km) and 810 ms^{-1} (300 km), see Figure 7.12. The best correlation between the velocity from GPS and from the ASC observations occurred at 110 km (E region) for a velocity of 767 ms^{-1} .

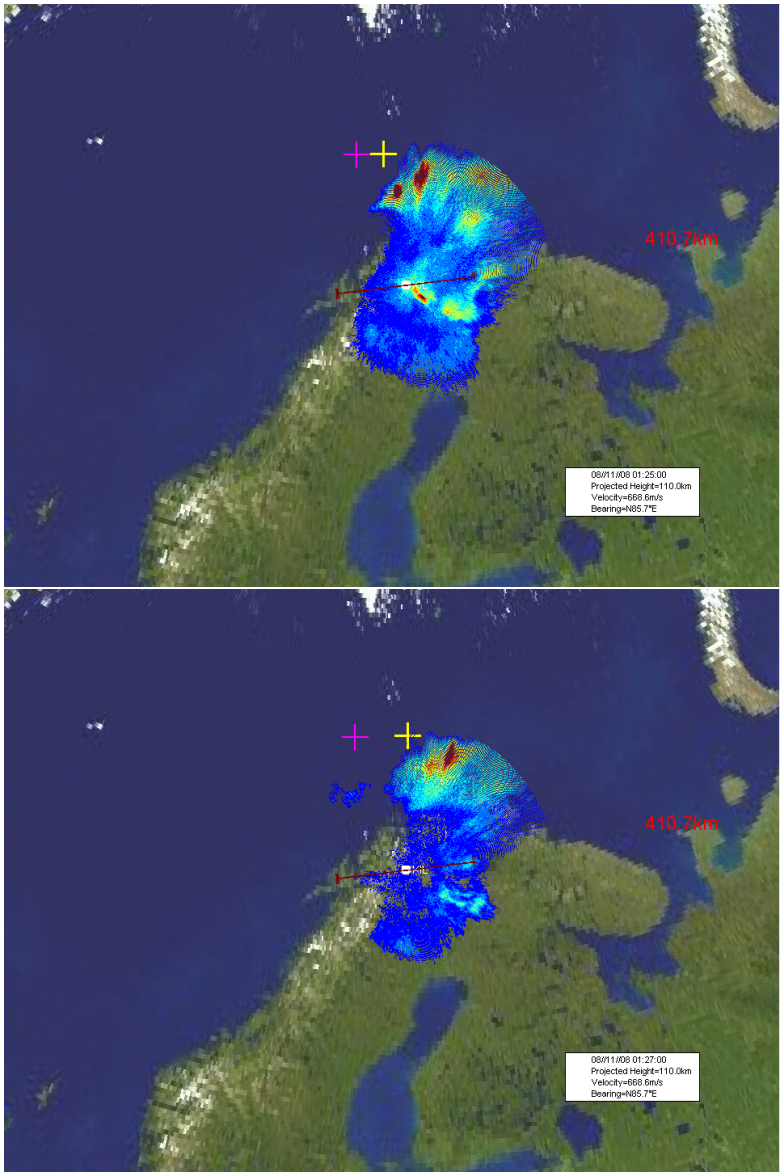


Figure 7.11: Kilpisjarvi ASC image at 01:25 and 1:27 UT on 8th November 2004, projected to 110 km altitude shell at approximately 2 and 4 minutes after the event. The yellow cross shows the projected location of the structure based on GPS receiver velocity measurements. The magenta cross shows where the receiver-satellite line intersects the altitude shell.

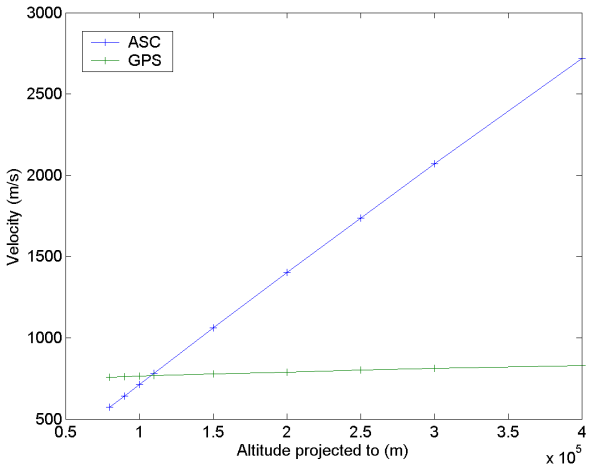


Figure 7.12: Measured Velocities from the all-sky camera images (blue) and the GPS receivers (green) for a range of ionospheric altitudes up to 400 km.

7.5 Simulating the Event

The auroral event has been simulated using the methods described in Chapter 4. The discriminator and bandwidth of the PLL and DLL are varied to determine how they affect the tracking of the signal.

Figure 7.13 shows the C/No of the received signal from the simulation using real measured scintillation and model generated scintillation. The left graphs show the results for narrow bandwidth tracking loops where as the right graphs show the results for wider bandwidth tracking loops. The results from the real amplitude and phase scintillation show good tracking for all but the standard Costas loop. The Costas suffers with loss of lock as soon as any significant scintillation is observed. The wider bandwidth version of the Costas does recover when the scintillation reduces. The Costas loop is particularly susceptible to amplitude fluctuations as the discriminator multiplies the amplitude values to obtain the error. This causes the error slope used to track the signal to vary with the square of the received signal power. Reducing the PLL bandwidth further makes the Costas loop more stable but also reduces the response. With the amplitude only scintillation, the narrow bandwidth Costas loop maintains a lock during all but the most severe scintillation, where as the wide bandwidth Costas is much less stable. The scintillation generated from one-minute averaged S_4 and σ_ϕ values show that the sharp fluctuations are damped out so no receivers lose lock. The WBMOD generated scintillation is even less variable, which can be explained as the model cannot predict the auroral structures, which are believed to be the cause.

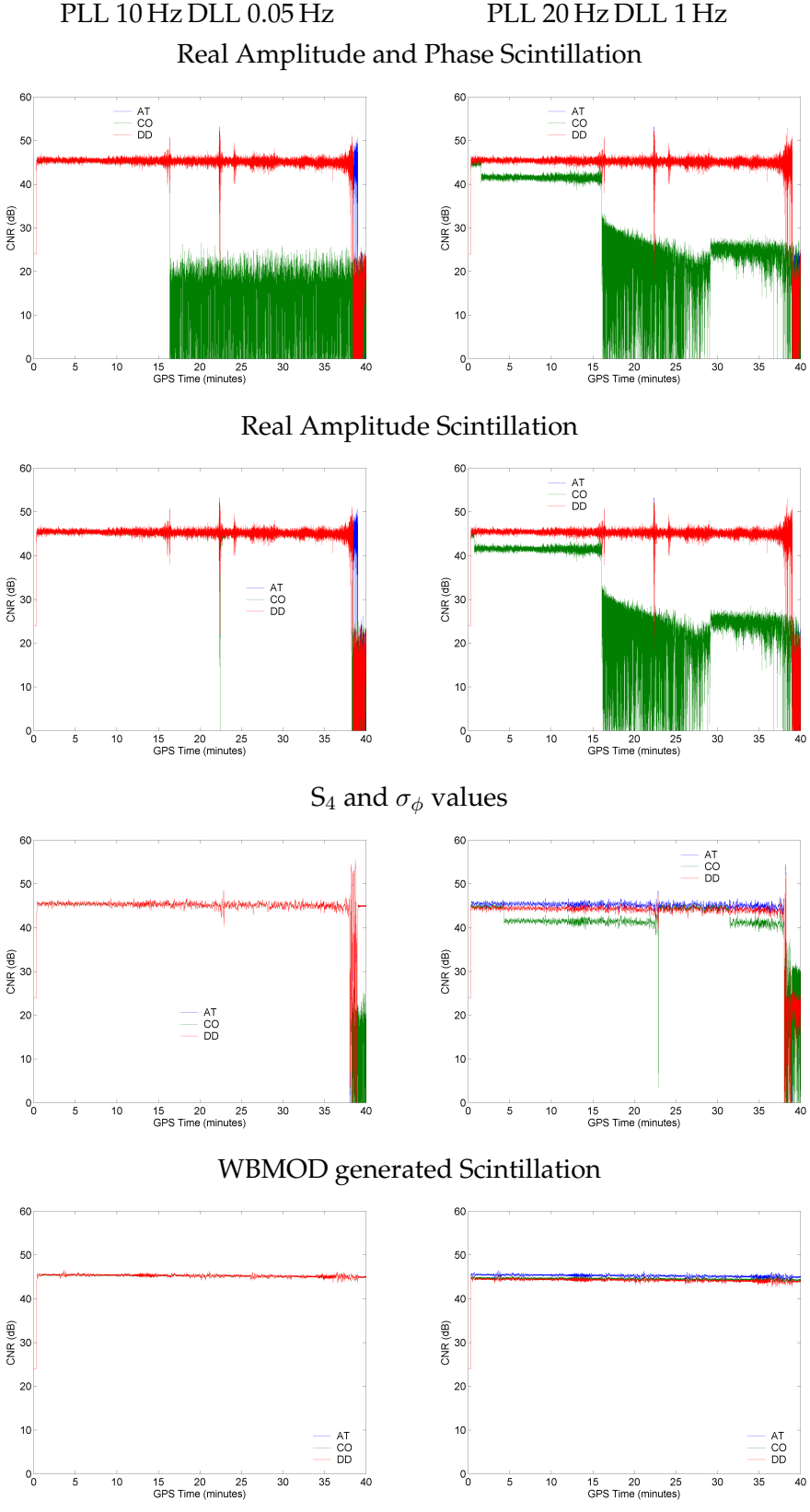


Figure 7.13: Shows the simulated receiver output using different discriminators and loop bandwidth. Discriminator key: AT = Arc-tangent, CO = Costas, DD = Decision directed

Figure 7.14 shows the positioning errors of the carrier (top) and code (bottom) tracking loops from the real scintillation simulation. The integrated carrier phase error from the arc tangent and decision directed trackers shows cycle slips during the peak scintillation point. The narrow bandwidth Costas tracker shows an accurate track until it losses lock where as the wide bandwidth Costas shows large errors from the beginning. The pseudorange shows small errors for all but the Costas PLL, which performs poorly in most of the cases.

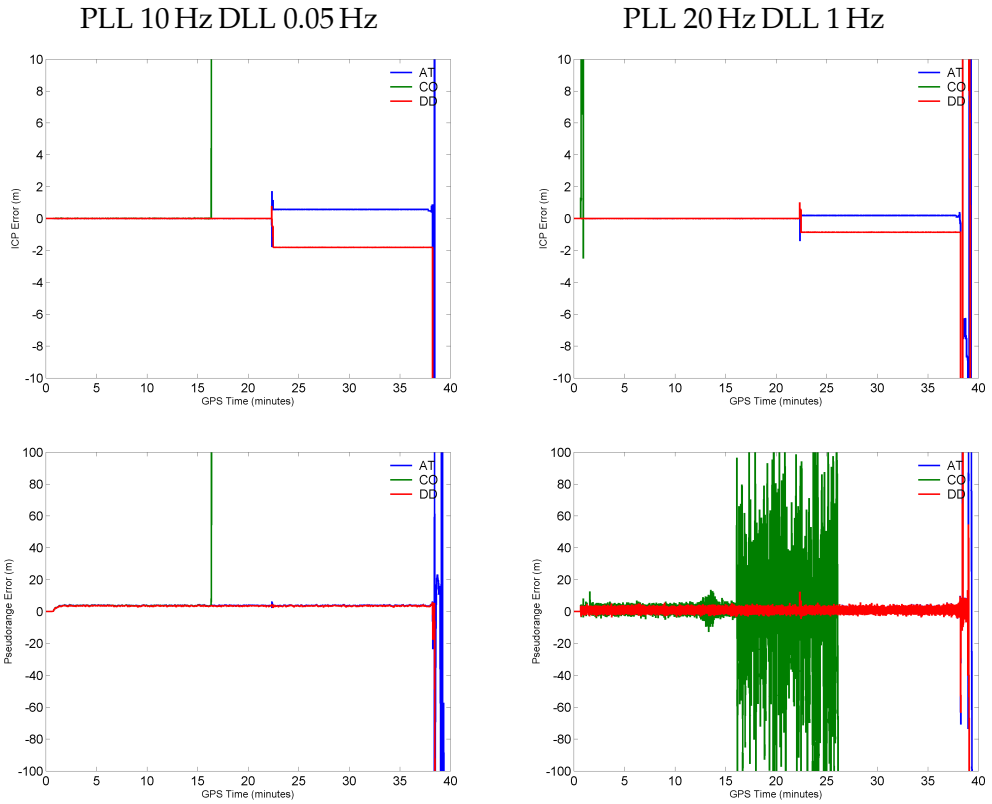


Figure 7.14: Shows the simulated receiver pseudorange and ICP error for various PLL bandwidths (i) and DLL bandwidth (ii)

Figure 7.15 shows the positioning errors from the simulation using only the real amplitude scintillation. The carrier-tracking error shows a similar performance to the simulation using amplitude and phase scintillation. This implies that the PLL is not significantly affected by the phase scintillation, only the amplitude scintillation in this case.

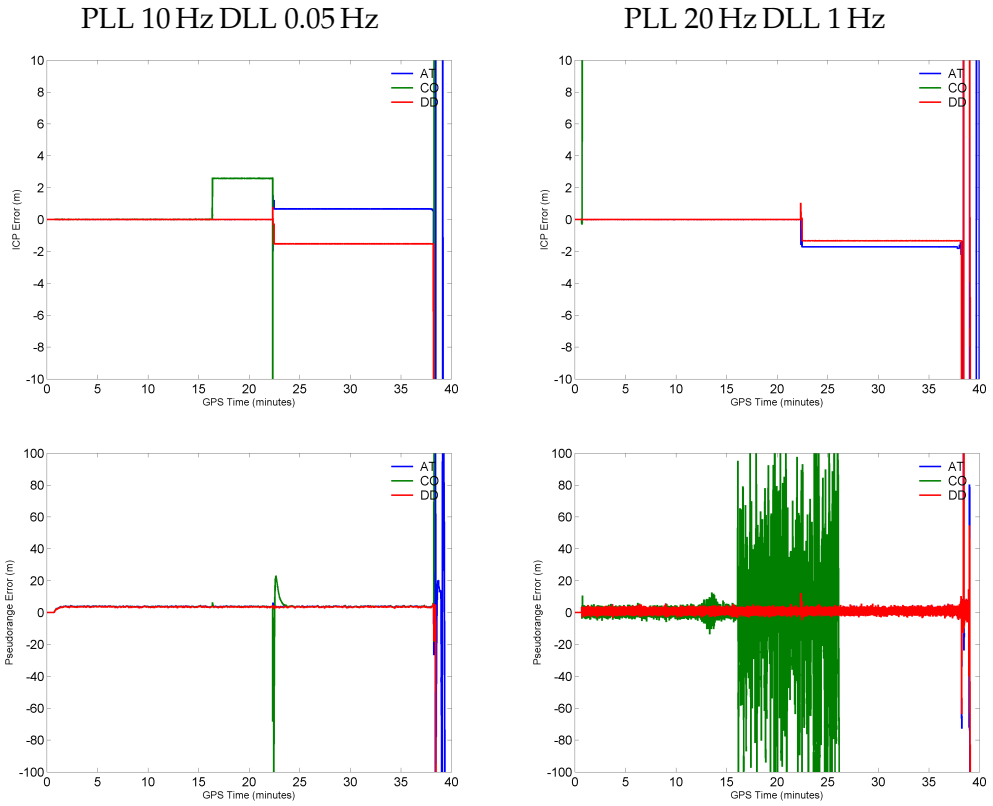


Figure 7.15: Shows the simulated receiver pseudorange error for various PLL bandwidths (i) and DLL bandwidth (ii)

Figure 7.16 shows the positioning errors from the simulation using the S_4 and σ_ϕ values. The carrier-tracking error shows good performance for all the narrow bandwidth trackers. Only the arc-tangent tracker performs well in all the experiments. In the wide bandwidth trackers both the Costas and DD ICP error are very large, however the pseudorange errors are small. The Costas and DD trackers have slightly lower C/N_0 values than the arc-tangent tracker. The wide bandwidth loops retain too much noise for the costas and DD trackers to form an accurate error slope to track with. The Costas tracker squares the noise during its I $\times$ Q operation which causes the large errors. The DD tracker effectively removes the noise from one channel by using a boolean operation, Q $\times$ sign(I) (I>0=1 I<0=-1), however if the I channel is noisy this becomes a random square wave which is less stable than the Costas method.

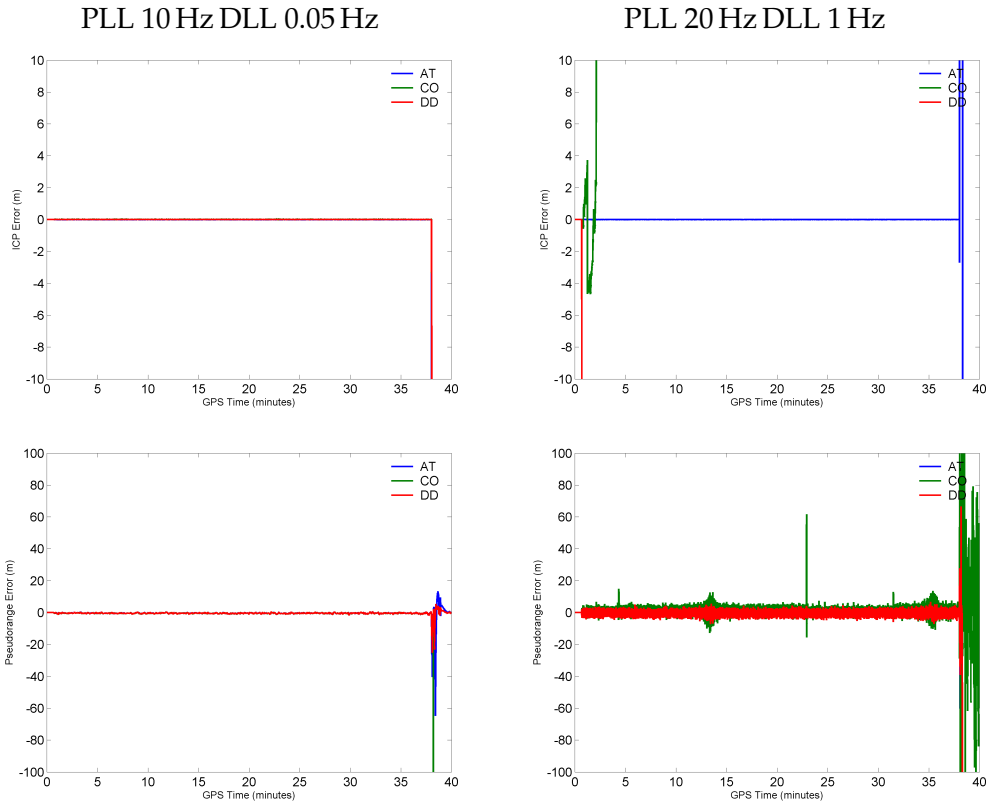


Figure 7.16: Shows the simulated receiver pseudorange error for various PLL bandwidths (i) and DLL bandwidth (ii)

The WBMOD simulation showed accurate code and carrier-tracking using narrow and wide bandwidth loops. This shows that the WBMOD model failed to accurately simulate the scintillation conditions, however it is not unexpected as any climatological model would fail to predict the small-scale fade investigated here.

7.6 Conclusion

Three closely located GPS receivers in the high arctic experienced a short period of rapid signal fading during a substorm event. Two of the receivers are from different manufacturers and are collocated; these showed almost identical fading patterns. Analysis using the MAT output and the GPS simulator showed that multi-path exists in the area, but confirmed that the rapid signal fade was not due to multi-path. Closer examination revealed a time difference between the fading events from the two receivers of the same type along a 240 m east-west baseline. This translated into a ionospheric event moving at between 758 ms^{-1} (80 km) and 810 ms^{-1} (300 km) eastwards. This corresponded well with the passage of a fast-moving auroral arc observed by all-sky camera data, showing evidence of particle precipitation that could be causing ionospheric

irregularities affecting the GPS signals. By comparing a range of different assumed altitudes for the event a good correlation between the all-sky camera and the GPS results was found at 110 km for a velocity of 767 ms^{-1} eastwards. The GPS scintillation is thus attributed to irregularities associated with a fast-moving auroral source of electron density precipitation.

The results indicate that the sudden fading seen here on the GPS L1 signal is likely to be related to ionospheric irregularities produced by auroral precipitation. This is the first study that shows a direct correspondence between the fading of GPS signal and the aurora; while this is a particularly intense substorm event it would be interesting in the future to use the fast-sampled GPS data in the auroral oval to investigate the frequency and duration of such occurrences. Techniques to automatically identify and investigate such features are under development (*Materassi and Mitchell*, 2007). The importance of such events to the reliable use of GPS for safety-critical applications is yet to be determined.

The GPS simulator was used to attempt to replicate the fade using the; real receiver data, S_4 and σ_ϕ values, WBMOD scintillation model. Obviously the real data provides the most accurate output, however the S_4 and σ_ϕ were used to evaluate their use as an scintillation metric. The results showed in this case that the S_4 and σ_ϕ could not provide the required information to accurately simulate the event as the fade was too short in duration. The effect of the real data on the positioning outputs from the simulator was much more pronounced than those from the S_4 and σ_ϕ. This shows that for these types of events S_4 and σ_ϕ are ineffective metrics.

7.7 Summary

This chapter has described a case-study investigating a rapidly changing but short duration signal fade observed on three co-located GPS receivers. The MAT and the GPS simulator were used to rule out multi-path as the cause of this fade. The ionosphere was investigated using a number of different measurement equipment based nearby the receivers. An auroral structure was identified in the region that the GPS signal propagated through and was determined to be the likely cause of the fade. The simulator was used with the scintillation measurements obtained from the real receivers to determine the effect of the rapid fade on different receiver architectures.

Chapter 8

Conclusions and Further Work

8.1 Conclusion

The aims of this research project were to:

- develop a simulation of the GPS which can reproduce the effects of the local environment and the ionosphere.
- investigate the effects that the local environment and the ionosphere has on the GPS signal.
- investigate what effect these have on the ability of the receiver to determine a position.
- determine criteria to distinguish between local environment and ionospheric errors from the receiver outputs.

A GPS signal simulator has been developed to analyse and model the effects of local multi-path and the ionosphere. No investigation of the ionospheric delay has been undertaken as its effect on positioning is well documented, however a delay bias based on the IRI model has been included in the simulator for future research. Ionospheric scintillation has been included as either a reference model (WBMOD) with noise distributions or by importing real receiver scintillation data. The multi-path model was developed using a computer graphics based ray-shooting algorithm providing accurate multi-path delays for reflected signals. Satellite hardware, receiver hardware and the signal generation have been modelled in the simulation. A number of experiments have been completed investigating the effects of the local multi-path and the ionosphere. The multi-path effects have been shown to be more predictable than ionospheric scintillation and a multi-path analysis method has been determined. A tool

has been developed that can analyse receiver power data and identify multi-path signatures. It can also provide some properties of where the multi-path arrived from, however interpretation of these values is currently difficult and requires some estimations. Experiments have been completed to investigate the effects of multi-path and scintillation on the receiver's ability to track the signals.

The first experiment described in Chapter 5, investigated the ability of the simulation to replicate the effects of a multi-path environment on a real receiver. A reflective panel was placed in the vicinity of the real receiver and the same reflective panel was modelled using the simulator. The results showed an oscillating interference pattern of similar phase and period in both the simulation and real receiver. The magnitude of the oscillating interference was more severe in the real receiver power output, which was attributed to differences between the real and simulated antenna polarisation rejection properties. The oscillating nature of the interference has been attributed to carrier interference, where the carrier signals of the direct and multi-path rays move in and out of phase. Another interference pattern was observed in the real receiver data with a much longer period than that of the reflective panel. By simulating the various objects surrounding the real receiver, the interference was found to be caused by the base of the weatherproof radome. This shows the ability of the simulator to be used as an environment analysis tool.

The experiment showed that carrier interference from multi-path can significantly affect the received power. Carrier interference was observed from both the radome base and the mounted reflector, however no cycle slips or loss of lock occurred throughout the experiments. This suggests a single ray from a nearby reflector does not cause any significant positioning errors in a receiver, as the multi-path signal is always a fraction of the direct signal and no (or little) code interference will occur. The magnitude of the carrier interference has been shown to be determined mainly by the polarisation rejection ratio of the antenna and electrical properties of the local environment. No code interference was shown in this experiment as it would require the multi-path to travel a greater distance (longer than 200m). This would distort the DLL error slope, where the severity of the effect will depend on the DLL type. This experiment has shown that single signal multi-path is easy to identify in the receiver's power output due to its oscillating pattern. It has also shown that the sinusoidal distortion of the signal has a period related to the distance of the reflector from the receiver and the satellite elevation. This periodic feature has been shown to be key in identifying multi-path errors from other errors using the receiver outputs.

The first experiment described in Chapter 6 investigates the relationship between interference period and the reflector distance further. A number of simulations were processed modelling environments with one large vertical reflector at various distances from the receiver. An oscillating interference pattern was again seen in the receiver

power outputs, and confirmed the relation to reflector distance and satellite elevation. The interference still had little effect on the PLL, where the integrated carrier phase error showed a 10 cm difference at most. This shows the benefits of using the carrier tracking values, whether to improve the precision of the pseudorange measurements or to use as a relative positioning system (Chapter 3). The DLL was significantly effected where a similar oscillating error to the power output is seen. The magnitude of the error is shown to be related to the distance of the reflector, where nearby ($<$4 m) and far away ($>$150 m) reflections cause small errors but mid-range (4 m -150 m) reflections cause large errors. This is shown to be caused by the error slope of the early-minus-late DLL (see Figure 5.15). A multi-path mitigation algorithm, like those described in Chapter 3, would significantly reduce this error. Unfortunately, no multi-path mitigation algorithms have been implemented in the simulation yet. The DLL errors show that the receiver implemented performs equally to a real receiver early-minus-late DLL. *Dierendonck et al.* (1992) shows a performance analysis of the early-minus-late correlator against the narrow correlator version.

The results from this experiment led to a method to identify and analyse multi-path signatures using the power output of a receiver. The method is derived in Chapter 6 and consists of a modified DFT that searches the possible satellite elevations and reflector distances. The method is shown to clearly identify multi-path signatures in the receiver power outputs, however the interpretation of the elevation can be difficult is some circumstances. The method assumes that there is a 2D plane connecting the positions of the receiver, the reflection point and the satellite. Therefore the plane may not be vertical and so the elevation value becomes difficult to interpret. Further development of the method should provide clearer results, until that time, combining the Multi-path Analysis Transform (MAT) output with an examination of the environment can determine the multi-path sources. This problem is greatly reduced in the simulator by viewing the ray-shooter outputs, so a combination of simulating the environment and using the MAT should provide a solution. The MAT method has shown encouraging results on both the simulator and real receiver outputs.

The next experiment investigated more complex realistic environments where many multi-path rays occur. The main conclusion from these simulations is that while a single multi-path ray is both easily identifiable and has little effect on the receiver tracking, multiple rays can severely distort the signal power and cause large tracking errors. A single reflected ray has been shown to cause an oscillating interference pattern, however multiple rays are shown to cause much more erratic patterns. This makes identifying interference patterns, even with ray-shooter outputs, very difficult. The method described earlier to analyse multi-path effects using the receiver power output has been further tested in these simulations. The transform has been shown to separate the effects of different multi-path signatures and can determine how much of a deviation the reflected ray has undergone in comparison to the direct ray. As discussed,

full interpretation of the outputs from the multi-path analysis can be difficult, however in multiple ray situations it is preferable to the signal power itself. A brief test of the multi-path analysis transform on real receiver data was completed, identifying a nearby structure as a source of multi-path when the satellite was in certain positions. Further testing is required under a more controlled environment, where the number of multi-path sources can be limited.

A case-study was completed investigating a short period of rapid signal fading observed in the outputs of three closely located GPS receivers in the high arctic. Two of the receivers are from different manufacturers and are collocated; these showed almost identical fading patterns. Analysis using the multi-path analysis transform and the GPS simulator showed that multi-path exists in the area, but confirmed that the rapid signal fade was not due to multi-path. Closer examination revealed a time difference between the fading events from the two receivers of the same type along a 240 m east-west baseline. This translated into a ionospheric event moving at between 758 ms^{-1} (80 km) and 810 ms^{-1} (300 km) eastwards. This corresponded well with the passage of a fast-moving auroral arc observed by all-sky camera data, that could be causing ionospheric irregularities affecting the GPS signals. By comparing a range of different assumed altitudes for the event a good correlation between the all-sky camera and the GPS results was found at 110 km for a velocity of 767 ms^{-1} eastwards. The GPS scintillation is thus attributed to irregularities associated with a fast-moving auroral source of electron density precipitation. This is the first study that shows a direct correspondence between the fading of GPS signal and the aurora.

The GPS simulator was used to attempted to replicate the fade using; the real receiver data, the S_4 and σ_ϕ values and the WBMOD scintillation model. Obviously the real data provides the most accurate output, however the S_4 and σ_ϕ were used to evaluate their use as an scintillation metric. The results showed in this case that the S_4 and σ_ϕ could not provide there required information to accurately simulate the event as the fade was too short in duration. The effect of the real data on the positioning outputs from the simulator was much more pronounced than those from the S_4 and σ_ϕ.

The project has shown a number of successes:

- The development and testing of the GPS signal simulator.
- The validation of the multi-path model with some limitations.
- The identification of a multi-path signature and an initial version of a tool to search for them.
- The identification of a signal fade associated with a moving auroral structure.
- The development of a number of software receivers for use with the simulator's generated signal.

- The assessment of the software receiver algorithms against scintillation and multi-path.

This project has shown the advantages of using a GPS signal simulator especially in the area of multi-path analysis.

8.2 Future Work

The current ray-shooting algorithm has been shown to perform well in the experiments undertaken in this project. For the next step of larger, more complex environments the issues of wide ray-spacing and a limited triangular mesh environment will cause a problem. One solution considered during the project was to use the hardware 3-dimensional accelerators used in computer graphics cards to process the algorithm. This would allow very high density rays and large triangular mesh environments, which would enable accurate multi-path modelling even in complex environments. Some success was achieved in developing this method during the project, however the field was in its infancy and some problems could not be solved. This is no longer the case as a number of "development kits" are available for the common graphics architectures which allow none specific software to be processed by the graphics card. Although this solution would be a great improvement, the algorithm still cannot model diffraction of refraction of the signal. There is the option of developing a more comprehensive propagation model, like the PWE of Finite Difference - Time Domain (FD-TD) methods, however these also have problems and require substantial computer resources. These types of methods are a longer term aim, but could provide improved understanding of the multi-path problem. The next step would be to perform an investigation, to determine whether this and other areas of the simulation would benefit from the parallel processing ability of the computer 3-D graphics card.

There are a number of parts to the simulation which although have been tested, have not yet been investigated due to time restrictions. This provides the opportunity to investigate a number of areas without further simulation development. These include investigating the effects of :

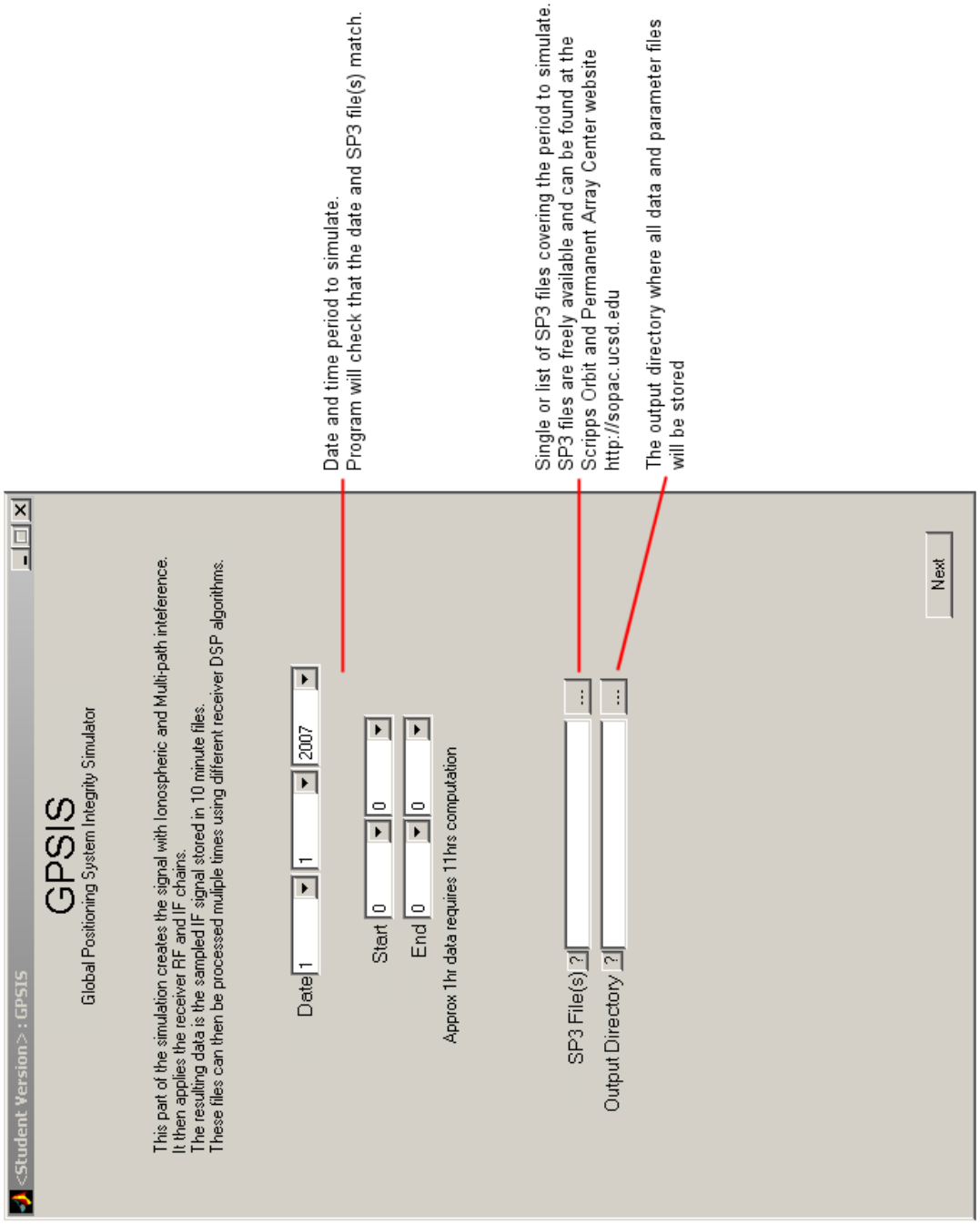
- receiver clock noise and drift.
- signal code distortion due to the receiver RF chain group delay or the ionosphere using the FFT.
- different sample rate and quantisation values.
- different correlation codes and code clock rates.

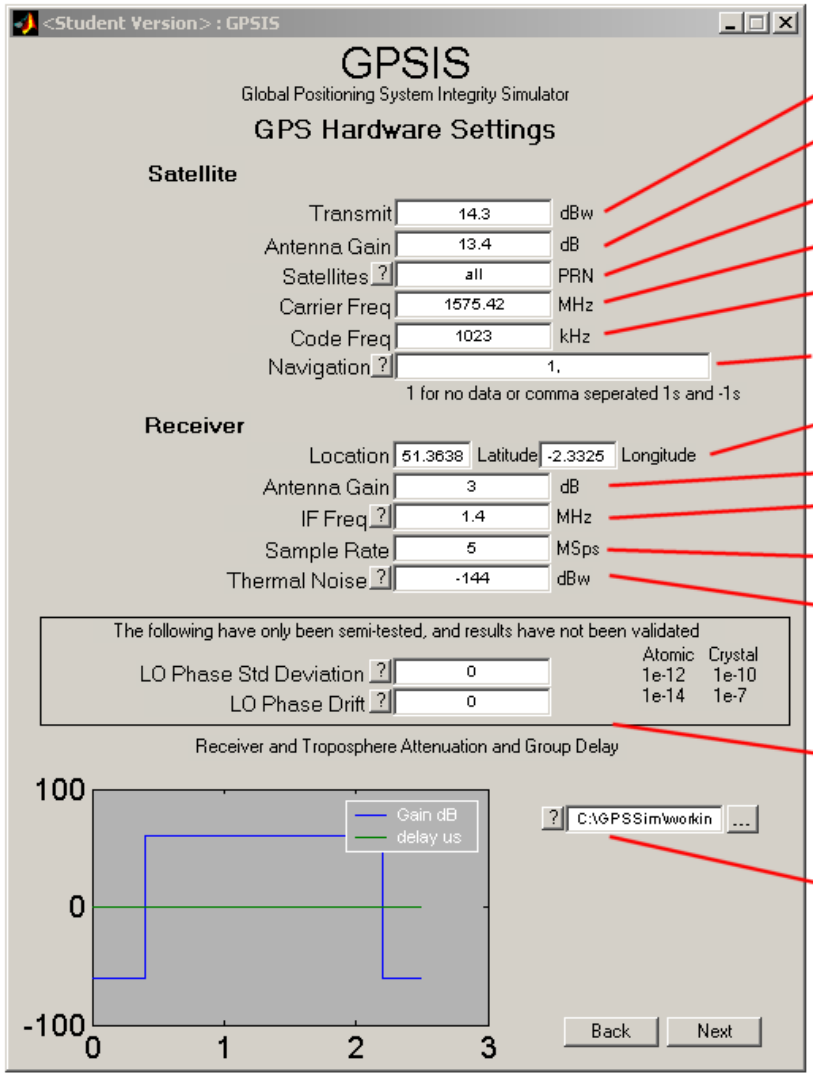
In terms of the signal generator development in the short term, the addition of the P code and the second frequency L2 would enable the evaluation of the PPS. In the longer term an investigation of the Galileo navigation system's signal structure and the GPS upgrades (described in Chapter 2) would be of great interest. On the receiver side, a number of improved algorithms (described in Chapter 3) can be implemented to evaluate their performance against multi-path and scintillation. The development of the narrow correlator DLL (*Dierendonck et al.*, 1992) and Van Nee's multi-path estimating algorithm (*van Nee*, 1992) would enable analysis and evaluation of the methods under controlled multi-path environments. The first stage of an ionospheric signal model has been developed, however this could be developed into a more realistic model like the PWE combined with random media approximations. This could provide a better the understanding of the ionosphere and its effects on the GPS system. A complete PWE model solution (from satellite to receiver) would be possible, but it would require significant computational resources. This would include; ionospheric diffraction and refraction, terrain reflection and diffraction, troposphere refraction.
With some further development the simulator could be used to analyse:

- Multi-path environments.
- Receiver architectures against multi-path and the ionosphere.
- Ionosphere effects on the GPS signal.

For the multi-path analysis transform (MAT) the problem of converting the values from those associated with a 2-dimensional plane of unknown orientation needs to be solved. With the measurements of the satellite elevation and azimuth, conversion to a more interpretable form should be possible.

Appendix A: Simulator User Interface





Satellite transmit power. GPS L1 power is default.

Satellite antenna gain. GPS L1 gain is default (no pattern currently).

Satellites to include. Default is all (simulator will determine which are in view)

Carrier frequency of signal. GPS L1 frequency is default.

Clock frequency of modulated code. GPS C/A frequency is default.

Navigation data. Default is no data, but simple bit patterns can be included.

Location of the receiver. Default is Bath University (N positive E positive)

Antenna gain of the receiver.

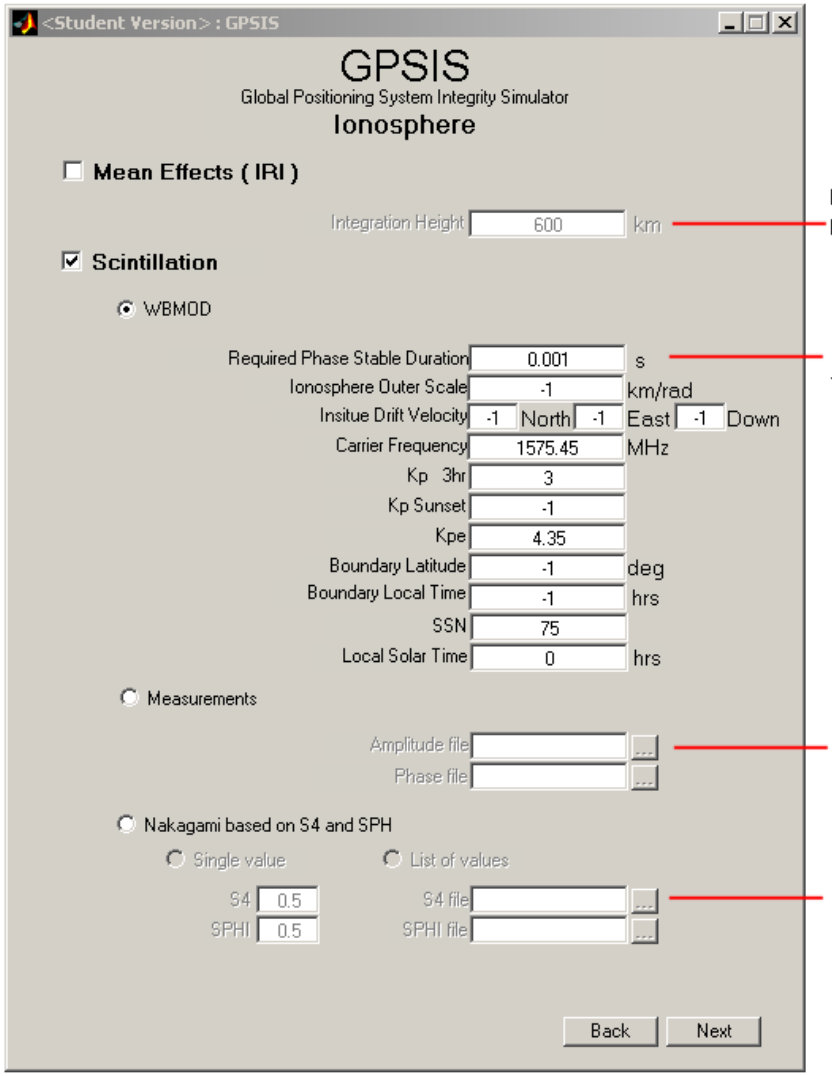
IF frequency to sample at

Sample rate of the signal. Must be at least double the IF frequency.

Thermal noise received by antenna. kTB (Chapter 4) T ~110K B~2MHz

Precision and accuracy of receiver clock. Default is no noise or drift.

IF filtering and group delay can be applied via files. Default is shown in graph (no delay, 1.8MHz bandwidth) Other more realistic files are available. Files consist of a list of 512 gain(dB) and delay (us) values.



IRI integration height. All other required parameters are taken from previous inputs

Standard parameters required by WBMOD. -1 values mean use WBMOD default.

If using real receiver values input filenames MATLAB workspace files with 50Hz measurements under the variable names 'ScintAmp' and 'ScintPha'. Detrending methods are described in Chapter 4

If using S4 and SPHI options are:
Single values for all satellites and times.
Single values of each satellite for all times.
Values for each satellite for each minute of time.
Multiple values can be applied using a file list

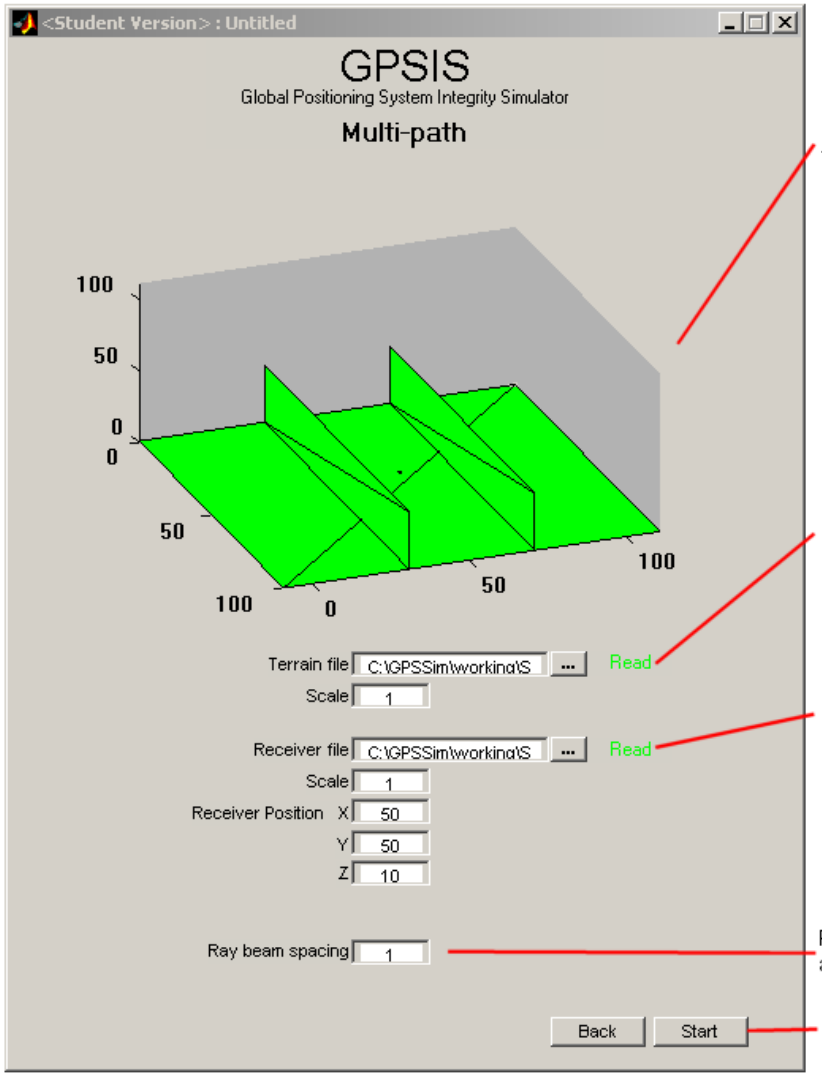


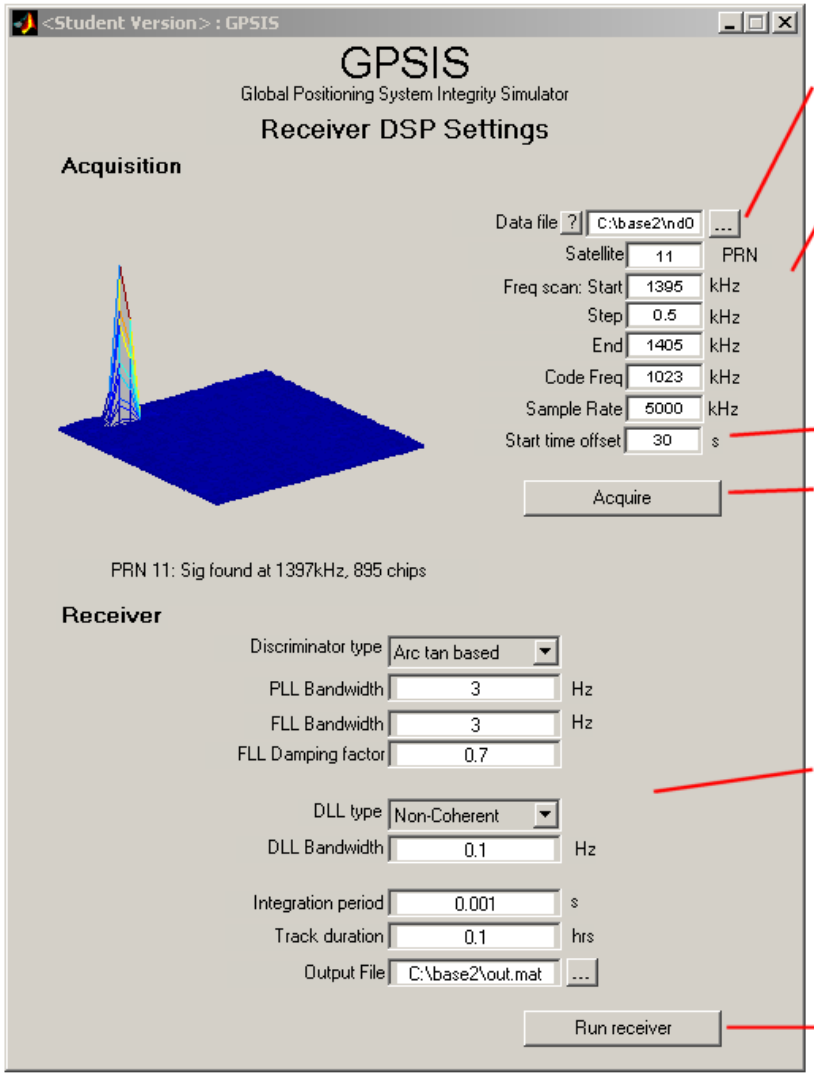
Diagram of the local environment defined by the terrain (green) and receiver (red) files.

File containing vertexes of each surface and their associated conductivity and permittivity
Scale determines the scale of the measurements in metres
A number of default environments are available.

File containing vertexes of each surface of the receiver.
-1 values for the conductivity and permittivity are used to determine if the vertex is part of the antenna or ground plane.
A number of default antennas are available
Position places the receiver in the environment.

Ray beam spacing determines how close the rays are to each other. Typically, smaller is better but slower.

Starts the simulation



Data file created by the signal generator.

Satellite identity and signal parameters required to acquire the signal.

Time offset before acquisition begins.

Acquire the signal.

Properties of the PLL, FLL and DLL.
Output is a mat file with an 'output' structure
The structure contains all the receiver outputs parameters

Run the signal tracker.

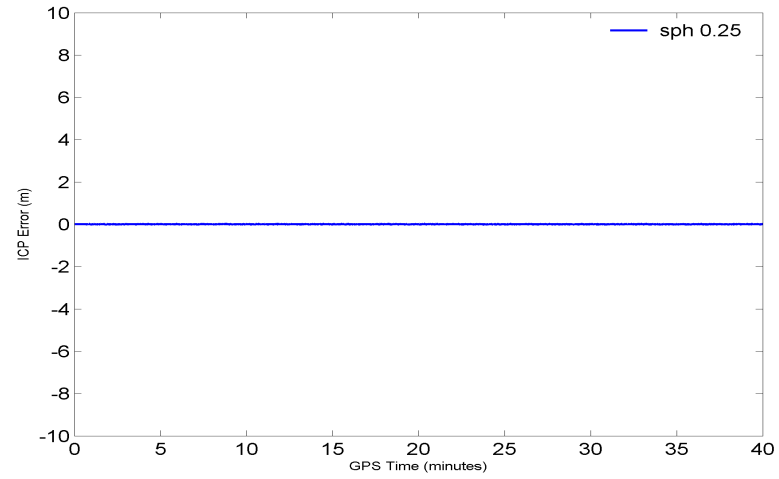
Appendix B: Effects of S4 and σ_ϕ combinations on receiver positioning

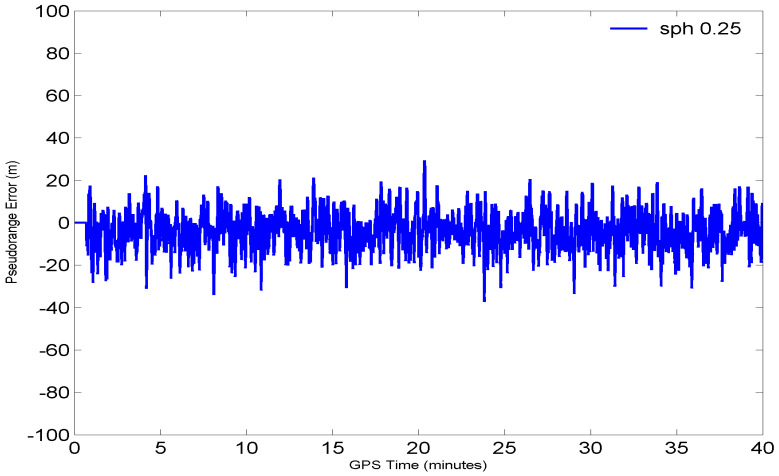
This appendix shows the effects combinations of amplitude and phase scintillation have on receiver positioning. The receiver specification is:

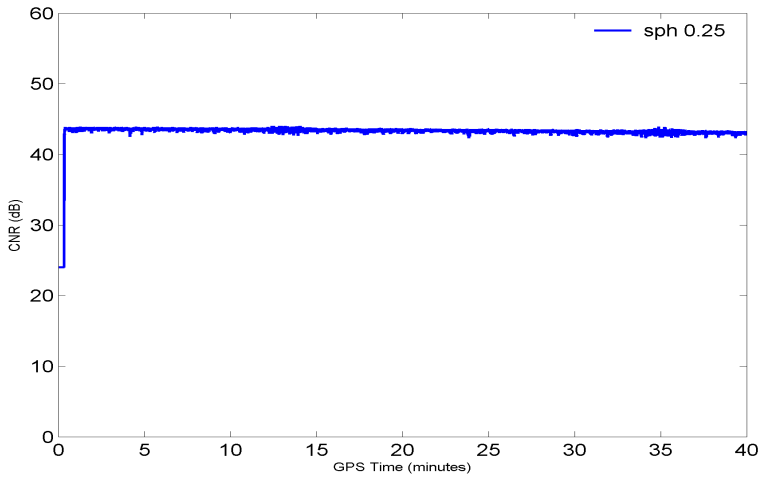
- Arc-tangent PLL with 10 Hz bandwidth.
- Arc-tangent FLL with 5 Hz bandwidth for initialisation and low signal strength tracking.
- Non-coherent early-minus-late DLL with 0.1 Hz bandwidth.
- Integrator period of 1 ms .

The results imply that the DLL is significantly more affected by the phase scintillation where as the PLL is more affected by the amplitude scintillation.

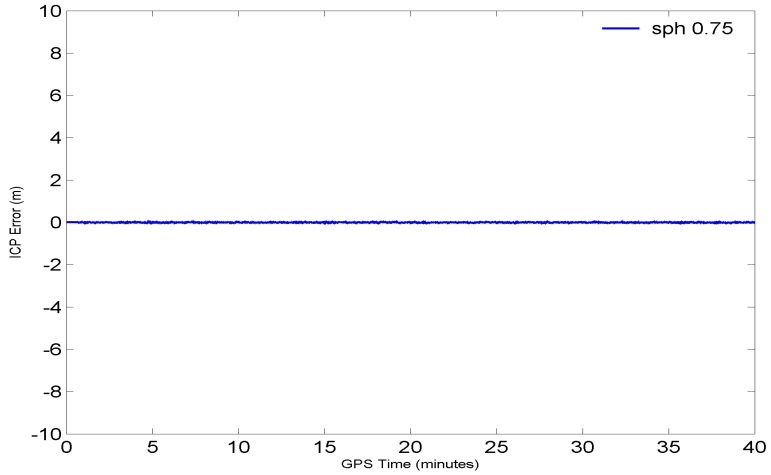
$\sigma_\phi = 0.25$

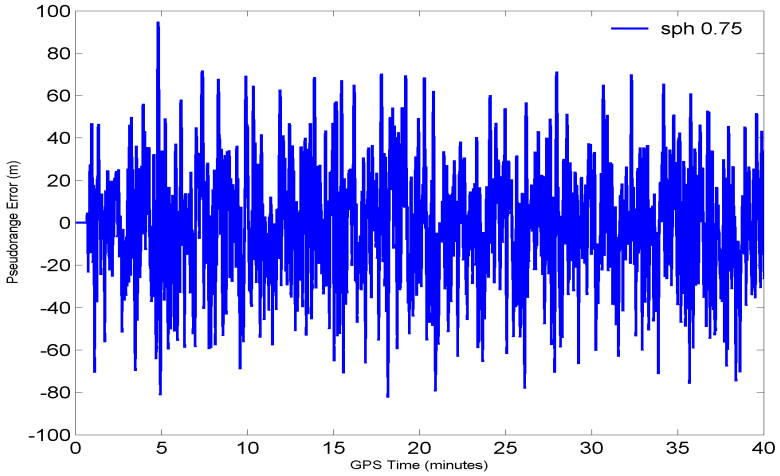


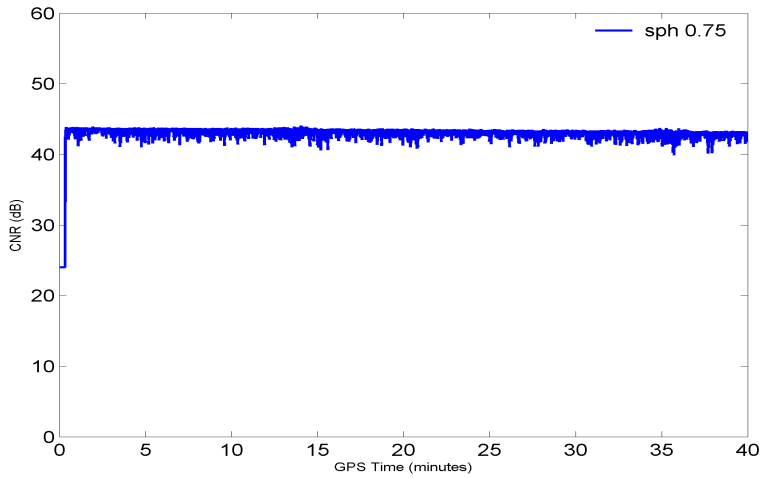




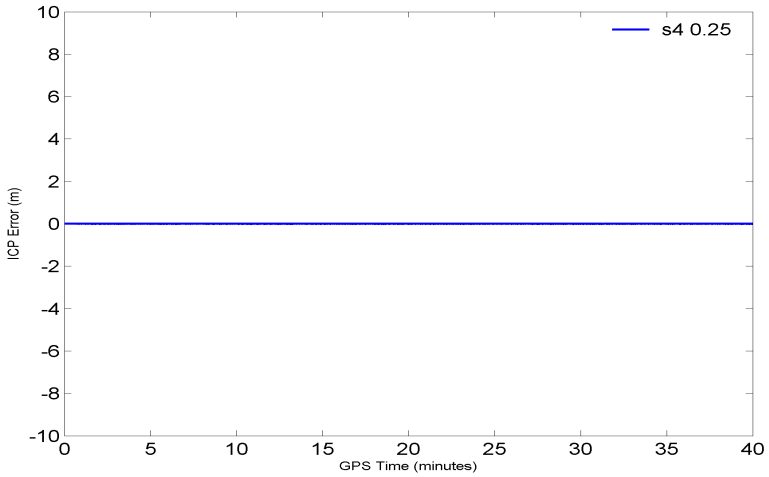
σ_ϕ=0.75

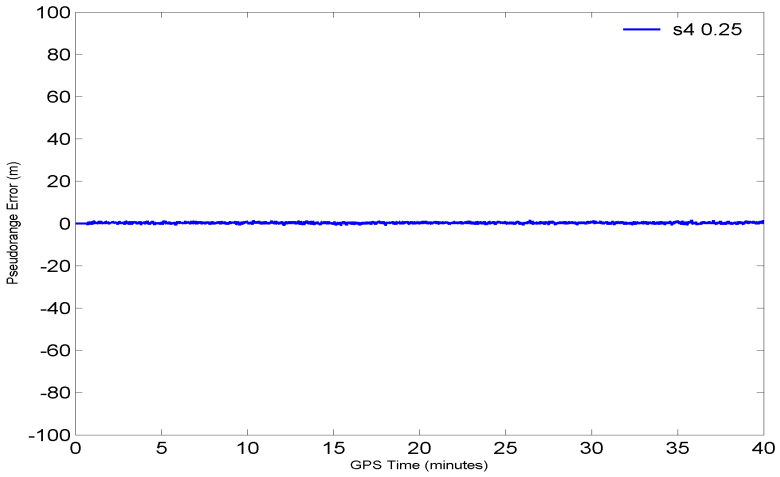


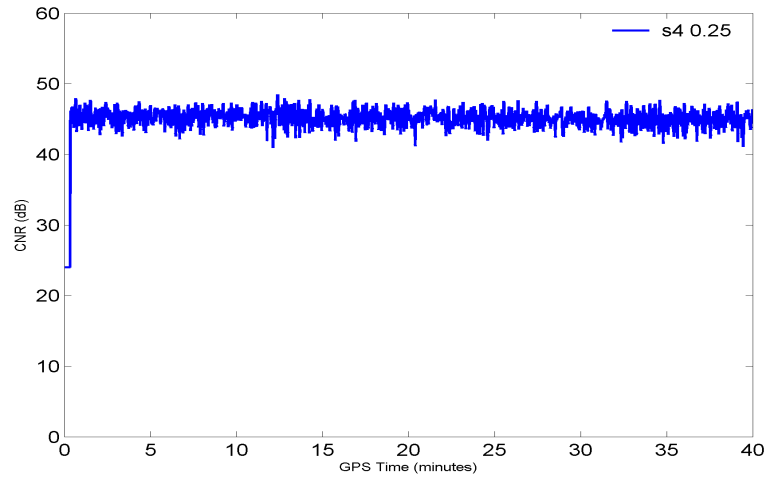




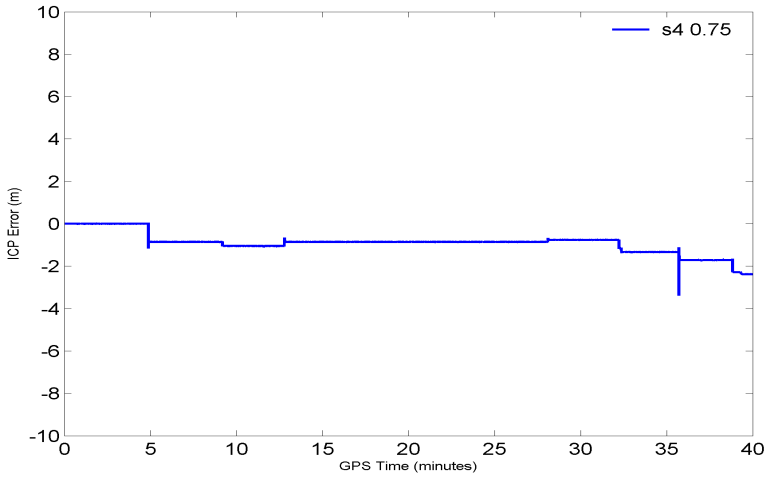
S_4=0.25

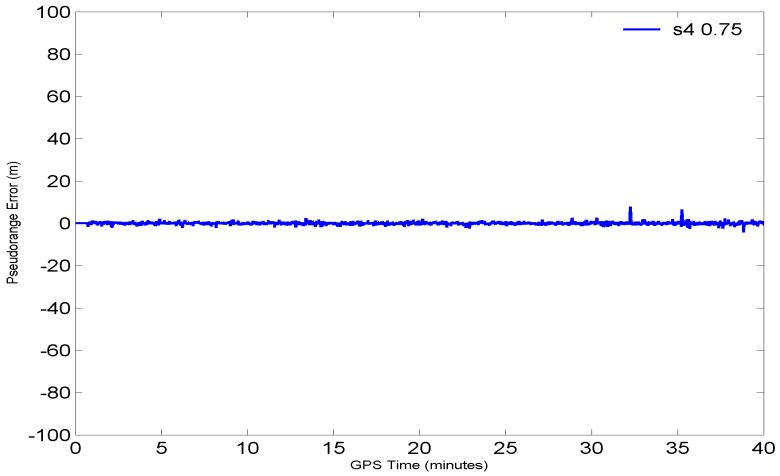

s4 0.25
ICP Error (m)
GPS Time (minutes)

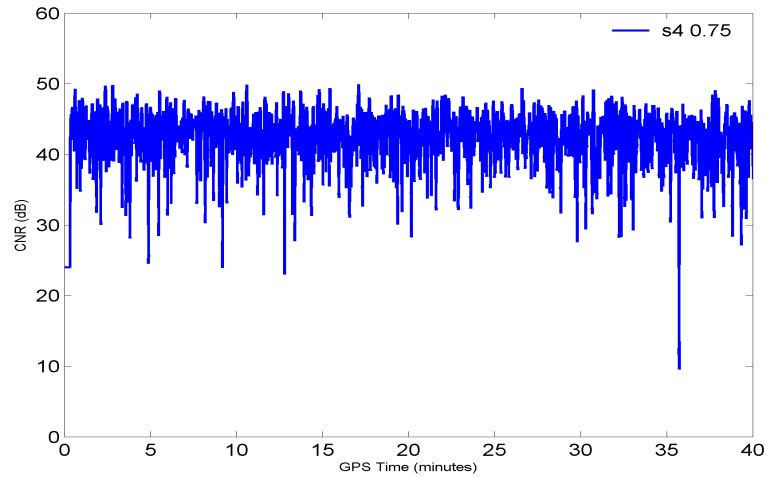

s4 0.25
Pseudorange Error (m)
GPS Time (minutes)


s4 0.25
CNR (dB)
GPS Time (minutes)

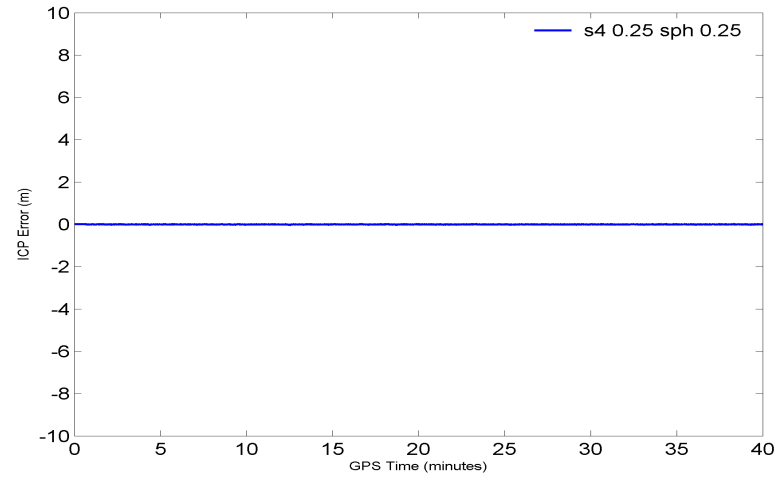
S_4=0.75

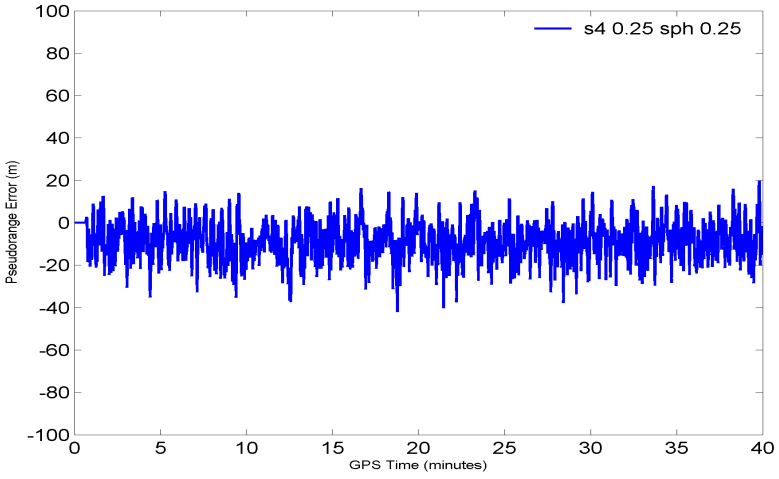

s4 0.75
ICP Error (m)
GPS Time (minutes)

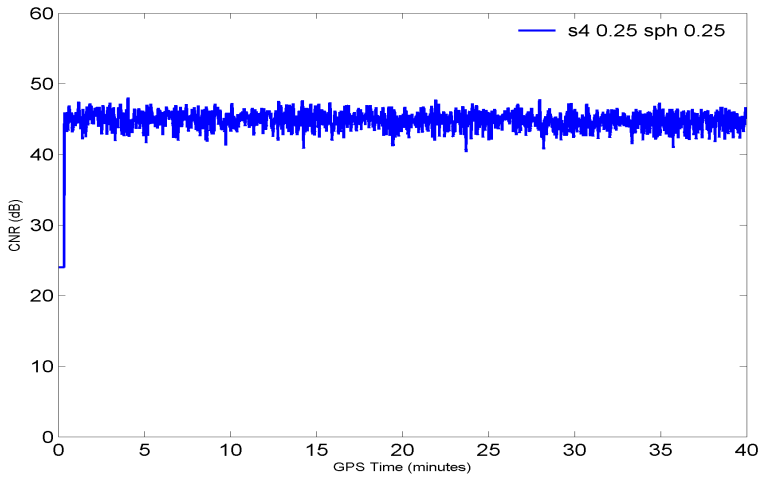

s4 0.75
Pseudorange Error (m)
GPS Time (minutes)


s4 0.75
CNR (dB)
GPS Time (minutes)

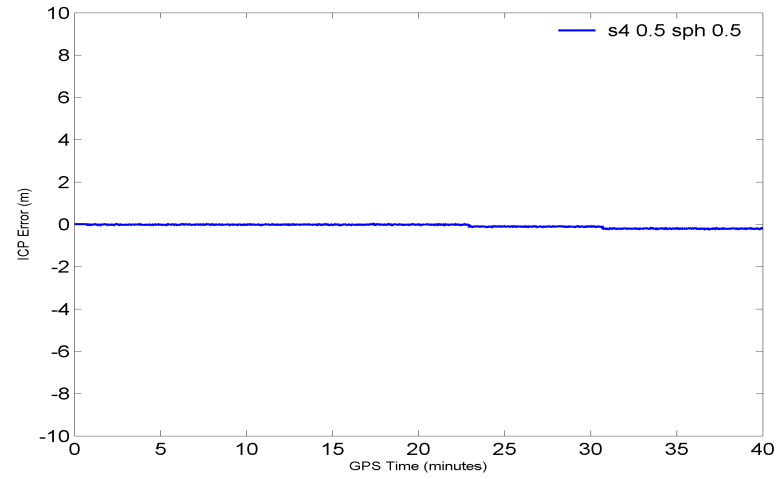
S_4=0.25 σ_ϕ=0.25

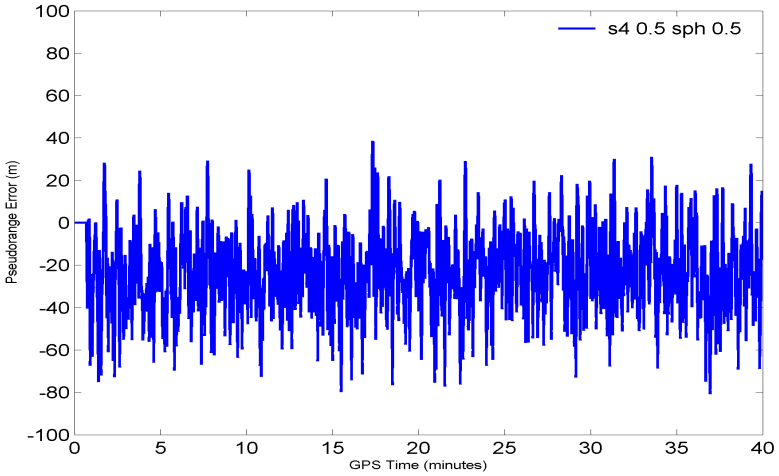


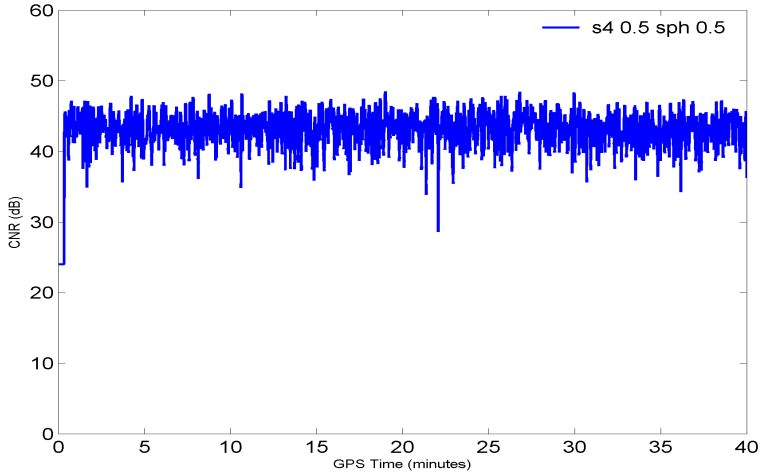




S_4=0.5 σ_ϕ=0.5







References

Basu, S., E. MacKenzie, and S. Basu, Ionospheric constraints on VHF/UHF communications links during solar maximum and minimum periods, *Radio Science*, *23*,3 , 363–378, 1988.

Basu, S., et al., Characteristics of plasma structuring in the cusp/cleft region at svalbard, *Radio Science*, *33*,6 , 1885–1899, 1998.

Basu, S., et al., Ionospheric effects of major magnetic storms during the international space weather period of september and october 1999: GPS observations, VHF/UHF scintillations, and in situ density structures at middle and equatorial latitudes, *J. Geophys. Res.*, *106*,A12 , 30,389–30,414, 2001.

Beach, T., and P. Kintner, Development and use of a GPS ionospheric scintillation monitor, *IEEE Trans. Geosci. Remote Sensing*, *39*,5 , 918–928, 2001.

Bean, B., and G. Thayer, Models of the atmospheric radio refractive index, *Proceedings of the IRE*, *47*,5 , 740–755, 1959.

Bilitza, D., International reference ionosphere - status 1995/96, *Adv. Space Res.*, *20*,9 , 1751–1754, 1997.

Bilitza, D., International reference ionosphere - 2000, *Radio Science*, *36*,2 , 261–275, 2001.

Bilitza, D., K. Rawer, L. Bossy, and T. Gulyaeva, International reference ionosphere - past, present, future, *Adv. Space Res.*, *13*,3 , 3–23, 1993.

Blewitt, G., Basics of the GPS technique: Observation equations., *Tech. rep.*, University of Newcastle, 1997.

Braasch, M., Performance comparison of multipath mitigating receiver architectures., *IEEE Proceedings of Aerospace Conference, 2001.*, *3*, 1309–1315, 2001.

Braasch, M. S., and A. J. V. Dierendonck, GPS receiver architectures and measurements., *Proceedings of the IEEE*, *87*,1 , 48, 1999.

Brent, R. I., and J. F. A. Ormsby, Scalar electromagnetic propagation modelling using parabolic equations and the split-step pade approximation., *IOP Journal of Physics A Mathematics and General*, *28*, 2065, 1995.

Brown, A., N. Gerein, and L. Savage, Multipath characterization using digital phased arrays., in *Proceeding of the Institute of Navigation 57th annual meeting*, p. 1, 2001.

Buchanan, W. J., Analysis of electromagnetic wave propagation using the 3D finite-difference time-domain method with parallel processing., Ph.D. thesis, Napier University, Edinburgh, 1996.

Burke, G., A physical optics model for scattering by irregular terrain at High Frequency (3 MHz-30 MHz) (HF), *Tech. rep.*, Lawrence Livermore National Lab., CA, United States, 1994.

Burklioltler, R. J., Fast physical optics integration for rough surface scattering., in *Antennas & Propagation Society Intl Symposium and URSI Natl Radio Science Meeting, IEEE*, 2003.

Chang, N., Presampling filtering, sampling and quantisation effects on digital matched filter performance., in *Proceedings of the International Telemetering Conference*, pp. 889–915, 1977.

Chen, H., L. Wang, and F. Chang, Long-duration carrier-smoothed-code algorithm for GPS positioning, in *Position Location and Navigation Symposium, IEEE 2000*, pp. 112–117, 2000.

Dai, D., T. Walter, C. Comp, Y. Tsai, P. Ko, P. Enge, and J. D. Powell, High integrity multipath mitigation techniques for ground reference stations, in *13th Canadian Symposium on Navigation, Ottawa*, 2000.

Dierendonck, A. V., P. Fenton, and T. Ford, Theory and performance of narrow correlator spacing in a GPS receiver., *Journal of the Institute of Navigation*, *39*,3 , 265, 1992.

Dierendonck, A. V., J. Klobuchar, and Q. Hua, Ionospheric scintillation monitoring using commercial single frequency C/A code receivers, in *Proceedings of ION GPS*, pp. 1333–1342, 1993.

Djebouri, D., A. Djebbari, and M. Djebbouri, Averaging correlation for fast GPS satellite signal acquisition in multipath Rayleigh fading channels., *Microwave Journal*, p. 66, 2004.

Doherty, P. H., S. H. Delay, C. E. Valladares, and J. A. Klobuchar, Ionospheric scintillation effects in the equatorial and auroral regions, in *Proceedings of ION GPS 2000*, 2000.

Essex, E., R. Birsa, N. Shilo, R. Thomas, M. Cervera, and A. Breed, Scintillation effects on global positioning system signals under solar maximum conditions, in *Proc. of the 2001 International Beacon Satellite Symposium, Boston, MA.*, 2001.

Frank, G. B., and M. D. Yakos, Next generation digital GPS receiver., *IEEE AES Magazine*, p. 10, 2000.

Fu, W., S. Han, C. Rizos, M. Knight, and A. Finn, Real-time ionospheric scintillation monitoring using GPS, in *Proc. of 12th Int. Tech. Meeting of the Satellite Division of the U.S.*, 1999.

Gustafsson, F., F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, and P.-J. Nordlund, Multi-configuration Kalman filter design for high-performance GPS navigation, *IEEE Transactions on Signal Processing*, *50*,2 , 425–437, 2002.

Hannah, B., Modelling and simulation of GPS multipath., Ph.D. thesis, Queensland University of technology, 2001.

Humphreys, T. E., M. L. Psiaki, P. M. K. Jr., and B. M. Ledvina, GPS carrier tracking loop performance in the presence of ionospheric scintillations, in *Proceedings of the 18th International Technical Meeting of the ION GNSS 2005, Long Beach, California*, pp. 156–167, 2005.

Hwang, P., and R. Brown, GPS navigation: combining pseudorange with continuous carrier phase using a Kalman filter, *Navigation: Journal of The Institute of Navigation*, *37*,2 , 181–196, 1990.

Hyaric, A. Z.-L., Wide-angle non-local boundary conditions for the parabolic wave equation., *IEEE Transactions on Antennas and Propagation*, *49*,6 , 916, 2001.

Kao, M., and D. Eller, Multi-configuration Kalman filter design for high-performance GPS navigation, *IEEE Transactions on Automatic Control*, *28*,3 , 304–314, 1983.

Kaplan, E., *Understanding GPS: Principles and Applications*, Artech House Books, 1996.

Kauristie, K., T. Pulkkinen, R. Pellinen, and H. Opgenoorth, What can we tell about global auroral-electrojet activity from a single meridional magnetometer chain?, *Ann. Geophys.*, *14*, 1177–1185, 1996.

Kraus, J., and D. Fleisch, *Electromagnetics with applications.*, McGraw-Hill, 1999.

Kroupa, V., *Phase Lock Loops and Frequency Synthesis.*, John Wiley & Sons, Ltd, 2003.

Kuttler, J. R., Differences between the narrow-angle and wide-angle propagators in the split-step fourier solution of the parabolic wave equation., *IEEE Transactions on Antennas and Propagation*, *47*,7 , 1131, 1999.

Ledvina, B. M., J. J. Makela, and P. M. Kintner, First observations of intense GPS L1 amplitude scintillations at midlatitude, *Geophys. Res. Lett.*, *29*,14 , 2002.

Lee, F. K. W., Modelling of electromagnetic wave propagation in printed circuit board and related structures., Ph.D. thesis, Multimedia University, Malaysia, 2003.

Lee, H., and C. Rizos, A new recursive carrier-smoothed-code filter for differential global navigation satellite systems, *11th Int. Assoc. of Institutes of Navigation World Congress, Berlin*, *263*, 2003.

Levy, M., *Parabolic equation methods for electromagnetic wave propagation.*, IEE Electromagnetic waves series 45, 2000.

MacDoran, P., R. Miller, L. Buennagel, and J. Whitcomb, Code-less system for positioning with NAVSTAR-GPS, in *Proceedings of the 1std International Symposium on Precise Positioning with the Global Positioning System*, vol. 1, 1985.

Materassi, M., and C. Mitchell, Wavelet analysis of GPS amplitude scintillation: A case study, *RADIO SCIENCE*, *42*,1 , 2007.

Galileo signal specification, *Galileo Open Service: Signal In Space Interface Control Document*, European Space Agency, 2006.

GPS performance document, *Global positioning system standard positioning service performance standard*, Office of the Secretary of Defence, 2001.

GPS signal specification, *Global positioning system standard positioning service signal specification*, USCG Navigation Center, 1995.

McCaskill, T. B., W. G. Reid, O. J. Oaks, R. L. Beard, and J. A. Buisson, Performance of global positioning system on-orbit NAVSTAR clocks., *IEEE International Frequency Control Symposium*, *49*, 133, 1995.

Miller, G. M., *Modern Electronic Communication*, Prentice-Hall, 1996.

Mitchell, C. N., L. Alfonsi, G. D. Franceschi, M. Lester, V. Romano, and A. W. Wernik, GPS TEC and scintillation measurements from the polar ionosphere during the october 2003 storm, *Geophys. Res. Lett.*, *32*,12 , 2005.

Moon, Y., and S. Verhagen, Integer ambiguity estimation and validation in attitude determination environments, in *ION GNSS 19th International Technical Meeting of the Satellite Division*, pp. 335–344, 2006.

Nakagami, M., *The m-distribution A general formula of intensity distribution of rapid fading in Statistical Methods in Radio Propagation*, 3-36 pp., ed. W. C. Hoffman, Pergamon Press, 1960.

Parkinson, B. W., and J. J. Spiker, *Global Positioning System: Theory and Applications Volume 1.*, American Institute of Aeronautics and Astronautics, Inc, 1996.

Rino, C. L., A power law phase screen model for ionospheric scintillation. I - weak scatter. II - strong scatter, *Radio Science*, *6*, 1135–1145, 1979.

Secan, J. A., R. M. Bussey, E. J. Fremouw, and S. Basu, An improved model of equatorial scintillation, *Radio Science*, *30*, 607–617, 1995.

Secan, J. A., R. M. Bussey, E. J. Fremouw, and S. Basu, High-latitude upgrade to the wideband ionospheric scintillation model, *Radio Science*, *32*, 1567–1574, 1997.

Teunissen, P., and S. Verhagen, GNSS phase ambiguity validation: A review, *Proceedings of SANE2007 Symposium,(IEICE), Japan*, *107*,2 , 1–6, 2007.

Townsend, B., J. Weibe, and A. Jakab, Results and analysis of using the MEDLL receiver as a multipath meter, in *ION National Technical Meeting*, 2000.

Turin, G., An introduction to digital matched filters, *IEEE*, *64*,7 , 1092–1112, 1976.

Umul, Y. Z., Modified theory of physical optics., *Optics Express*, *12*,20 , 4959, 2004.

Van Dierendonck, A., J. Klobuchar, and Q. Hua, Ionospheric scintillation monitoring using commercial single frequency C/A code receivers., in *Proceedings of ION GPS-93, Salt Lake City, UT*, pp. 1333–1342, 1993.

van Nee, R., The multipath estimating delay lock loop., in *IEEE 2nd International Symposium on Spread Spectrum Techniques and Applications*, 1992.

Verhagen, A., The GNSS integer ambiguities: estimation and validation., Ph.D. thesis, Delft Institute of Earth Observation and Space Systems, Delft University of Technology, 2005.

Viljanen, A., and L. Häkkinen, IMAGE magnetometer network. in satellite-ground based coordination source book (eds. M. Lockwood, M.N. Wild and H.J. Opgenoorth), *ESA publications SP-1198*, pp. 111–117, 1997.

Ward, P. W., GPS receiver search techniques., in *IEEE Position Location and Navigation Symposium*, p. 604, 1996.

Woo, K. T., Optimum semi code-less carrier phase tracking of L2, in *The 12th International Technical Meeting of the Satellite Division of the Institute of Navigation, Nashville*, 1999.

Yee, K. S., Numerical solution of initial boundary value problem involving maxwells equations in isotropic media, *IEEE Trans. Antennas and Propagation*, *14*, 302–307, 1966.

Zalin, D., and K. Sarahandi, Numerical simulation of scattering from rough surfaces using a fast far-field iterative physical optics approach., *IEEE Antennas and Propagation Society International Symposium*, *2*, 1228, 1999.